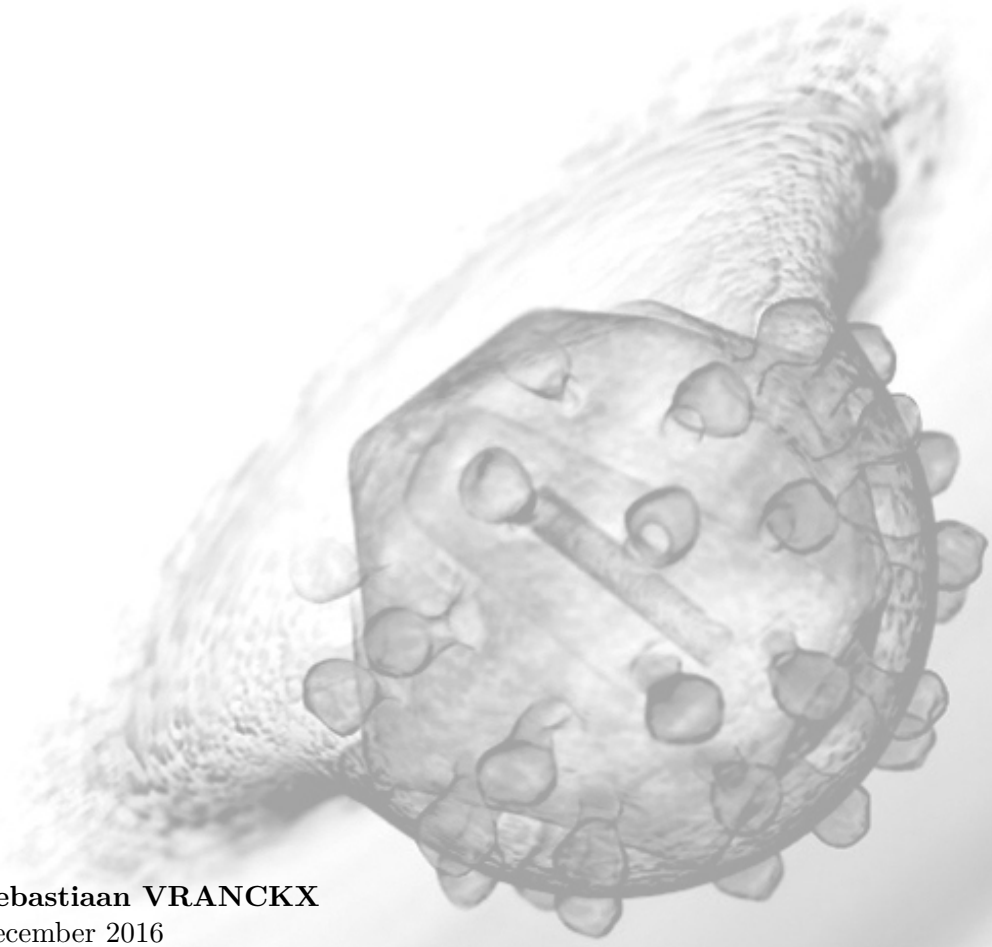
KU LEUVEN

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

FACULTY OF MEDICINE

IMPACT OF LEDGF/P75-IN INTERPLAY ON HIV INTEGRATION AND LATENCY



Lenard Sebastiaan VRANCKX
Leuven, December 2016

KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Pharmaceutical and Pharmacological Sciences
Laboratory for Molecular Virology and Gene Therapy

KU LEUVEN
DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

IMPACT OF LEDGF/P75-IN INTERPLAY ON HIV INTEGRATION AND LATENCY

Lenard Sebastiaan VRANCKX

Program:	MOLECULAR AND STEM CELL MEDICINE BIOMEDICAL SCIENCES
Promoter:	Prof. Dr. Zeger Debyser
Co-promoter:	Prof. Dr. Rik Gijsbers
Jury members:	Prof. Dr. Carine Van Lint
	Prof. Dr. Sharon R. Lewin
	Prof. Dr. Jan Cools (TAC)
	Prof. Dr. Kristel Van Laethem (TAC)
Academic year:	2016 - 2017 (Leuven)

Dissertation presented in partial fulfillment
of the requirements for the degree of
Doctor in Biomedical Sciences

Public Ph.D. Defense
Monday December 12th, 1 pm
AZC AUDITORIUM (UZRN 01.21)
Kapucijnenvoer 35 blok N
3000 Leuven

Uitgegeven in eigen beheer, Lenard Sebastiaan VRANCKX,
Kapucijnenvoer 33, VCTB +5, B7001 - 3000 Leuven (Belgium)

The cover depicts the lentiviral entry step (Illustration by Cortical Studios).

Acknowledgements

"BOUJAHHH", at last!

The process of writing this doctoral dissertation has been long and arduous. Little by little, a little became a lot. It has been over 9 years now since I embarked on a university degree, followed by a PhD program in the wonderful world of science, where off note time is a relative concept. Those decisions have led to one of the most exciting journeys one can imagine. Looking back, I've experienced several transitions (to say the least) both as a scientist and as a person. This infamous acknowledgements section denotes my appreciation and sincere gratitude to all friends, colleagues, family, etc. that supported and encouraged me throughout the years. Moreover, it allows me to apologize to those people that sometimes had to endure my complaints, experience my temper fluctuations or for whom I should have made some more time.

First and foremost, Zeger, I am truly thankful to have had you as my promoter. Allowing me to become part of the MolMedian community by first accepting me as a master student and next by stimulating me to embark on a PhD program has been a truly unforgettable experience. I admire your contagious enthusiasm for science and your capability to manage both a scientific and political career. Thank you for guiding my focus and allowing me to present my work at the CSHL meeting in NYC.

Rik, you have been one particularly great co-promoter! I admire your approach in science and thank you for all your training, guidance, time and unstinting support. In particular I appreciate the spontaneous meetings, diverse discussions and your critical/ creative input. I've come to hold you in high regard. You've been of paramount importance in helping me achieve this work. May we look forward to the many wine and whisky tastings that are yet to come.

My particular gratitude goes to all members of the examination committee: Prof. Carine Van Lint, Prof. Sharon R. Lewin, Prof. Kristel Van Laethem and Prof. Jan Cools. Thank you for your valuable input on consolidating and reviewing this thesis. Your remarks profoundly improved the quality and technical soundness of this manuscript.

Paulien and Babs, I've been extremely lucky to have worked with such incredible technicians. Thank you for all your help, support, friendship and laughter during the years!

Dear MolMedian labbies, thank you for all the wonderful years in the lab! I will forever cherish the crazy discussions, daily dose of fun, 'HIV lab'- gossip, BBQs, unforgettable lab weekends, laughter, dynamic atmosphere, ... irrespective of the unbalanced gender ratio. I will miss you all and the excessive amount of birthday treats!

TWITS family and friends, having me over for coffee during lunch was of major contribution to obtaining my PhD. "Het bankske" represented a moment of escape from the everyday working environment. The amount of brotherhood and the ledgendary Wednesday evening Eindhoven skate sessions are absolutely amazing. Thank you 'biochemistry possy' for the iteration of dinners, I much appreciate the continuous effort to keep this tradition alive. HSP, I will 'Never Forget' the incredible moments we had

and will have. PCV, thank you for all luminous days in paradise. You are awesome!

Mom - Dad: thank you, as cliché as it may sound, for giving me the opportunity to embark on a university degree and for supporting me in every possible way. Thank you for your unconditional love, confidence and for never doubting the decisions I made. Thank you for everything. Thank you for those things I sometimes took for granted or that went unnoticed. Thank you for your forgivingness and for caring when I was feeling a bit off. I am so fortunate to have such great parents, I love and admire you each in your own way. Also you Elise, Zekke - zusje, your strength, energy and positivity are unsurmountable and transitioned onto those in your close surroundings. Thank you for everything!

And last but far from least: Lauranne, Lorrie, liefje, kleinie, ... This dissertation manuscript would have never seen the light of day without you! Thanks for all your patience, love and unwavering support. Thank you for being my smile every single day! No words enough to describe how much I look forward to our future together and our not so little trip to the land of Oz. A new chapter, so ready, let's hit the road! 100 % LIEFDE!

And now, I believe, the well-deserved beer compels me to stop here.

Cheers, this one is for you!

Lenard, Le, Lele

Summary

A hallmark of retroviral replication is the reverse transcription of the viral RNA genome into a double stranded DNA copy that becomes stably integrated into the host chromosomal DNA by means of the viral integrase (IN). This integration event forever links the fate of the invading virus to that of its host cell. Retroviral integration is a non-random process and depending on the genus, integration occurs near specific features of the host genome. HIV-1 and other lentiviruses preferentially integrate into the body of actively transcribed genes. Much alike other viruses, HIV-1 relies on the presence of essential cellular co-factors to complete its viral replication cycle. In 2002 our lab identified Lens Epithelium-derived Growth Factor (LEDGF/p75) as a binding partner of HIV-1 IN [Cherepanov et al., 2003] responsible for orchestrating HIV-1 integration into active transcription units through interaction with the viral IN via its Integrase Binding Domain (IBD) [Ciuffi et al., 2005; Gijsbers et al., 2011a; Schrijvers et al., 2012b] and interaction with the host chromatin via its PWWP domain [De Rijck et al., 2010; Pradeepa et al., 2012; Eidahl et al., 2013; Gijsbers et al., 2011a; van Nuland et al., 2013]. Depletion of LEDGF/p75 shifts integration out of transcription units [Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007; Schrijvers et al., 2012b]. Additionally, replacement of the chromatin binding domains of LEDGF/p75 with alternative domains retargets integration near genomic loci bound by these domains [Ferris et al., 2010; Gijsbers et al., 2009].

In my Ph.D. project I continued along this research line and studied the mechanism of lentiviral integration and the role of LEDGF/p75 therein. The LEDGF/p75-IN interplay was studied to better understand the relationship between integration site selection, subnuclear positioning and proviral gene expression forming the common thread throughout the chapters. In this doctoral dissertation two research manuscripts are bundled.

In the first part of my work (see Chapter 3), I exploited chimeric versions of LEDGF/p75 as a tool to retarget lentiviral integration to potentially 'safer' regions within the human genome. As such, Serious Adverse Events (SAEs) associated with the use of retroviral vectors in gene therapeutic applications may be overcome. I designed LEDGF/p75-based tethers that resulted in a more random integration pattern which is generally believed to reduce the risk or probability of transcriptional dysregulation of neighbouring genes, an event dubbed 'insertional mutagenesis' [Chatziandreou et al., 2011; Derse et al., 2007; Kaufmann et al., 2013; Suerth et al., 2012].

Secondly (see Chapter 4), I studied the importance of integration site distribution in the context of HIV persistence. The development of antivirals and combination AntiRetroviral Therapy (cART) in the last decades revolutionized the treatment of HIV/AIDS, however, the major impediment towards a cure for HIV infection lies in the existence of long-lasting reservoirs containing transcriptionally silent but replication competent HIV provirus that is unaffected by cART [Trono et al., 2010]. From these

quiescent or latent reservoirs HIV rebounds within a few weeks after cessation of cART. Since integration is believed to affect the transcriptional activity of the integrated provirus, I assessed whether targeting and integration site selection affect the reactivation potential of the quiescent integrated provirus. Therefore I evaluated whether LEDGINs, *bona fide* small molecule inhibitors (2-(quinolin-3-yl)acetic acid derivates) binding to the IN dimer interface and thereby blocking the LEDGF/p75 interaction and interfering with the catalytic activity of IN [Christ et al., 2010, 2012], could affect the functional reservoir formation. I studied both (i) the establishment of HIV-1 latency and (ii) the reactivation from latency following treatment with different Latency Reversing Agents (LRAs).

Both studies on the role of the LEDGF/p75-IN interplay in integration site selection and HIV persistence either contribute to the development of safer viral vectors for gene therapeutic applications or to alternative approaches to cure HIV by aiming for a modulation or prevention of the generation of a residual HIV-1 reservoir able to rebound from latency once drug administration is withdrawn.

Samenvatting

Een typerend kenmerk van retrovirale replicatie is de omgekeerde transcriptie van een enkelstrengig viraal RNA genoom tot een dubbelstrengige DNA kopij, die vervolgens stabiel wordt geïnsereerd in het gastheerchromatine via het viraal integrase (IN). Dit integratie-evenement verbindt het lot van het provirus onomkeerbaar met dat van de cel. Ondertussen weten we dat deze insertie niet op een random manier gebeurt maar dat integratie, afhankelijk van het genus, welbepaalde voorkeursplaatsen in het genoom verkiest. Het Humaan ImmunodeficiëntieVirus (HIV-1) en andere lentivirussen integreren preferentieel in actieve genen. Net zoals alle virussen, is ook HIV-1 afhankelijk van de aanwezigheid van essentiële eiwitten in de cel om zijn replicatiecyclus te kunnen voltooien. In 2002 identificeerde onze onderzoeksgroep het Lens Epithelium-derived Growth Factor (LEDGF/p75) eiwit als een bindingspartner van HIV-1 IN [Cherepanov et al., 2003]. LEDGF/p75 is verantwoordelijk voor het sturen van HIV-1 integratie naar actieve transcriptie-eenheden door een interactie met het viraal IN via zijn Integrase Binding Domain (IBD) enerzijds [Ciuffi et al., 2005; Gijsbers et al., 2011a; Schrijvers et al., 2012b] en een interactie met het gastheerchromatine via zijn PWWP-domein anderzijds [De Rijck et al., 2010; Pradeepa et al., 2012; Eidahl et al., 2013; Gijsbers et al., 2011a; van Nuland et al., 2013]. Depletie van LEDGF/p75 in de cel stuurt integratie weg van transcriptie-eenheden [Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007; Schrijvers et al., 2012b]. Wanneer het chromatine-bindende domein van LEDGF/p75 wordt vervangen door alternatieve domeinen is er een preferentiële toename van integraties in de buurt van de loci herkend door deze domeinen [Ferris et al., 2010; Gijsbers et al., 2009].

Mijn PhD-project bouwt verder op deze onderzoekslijn. Ik bestudeerde het mechanisme van lentivirale integratie en de rol van LEDGF/p75 hierin. De wisselwerking tussen LEDGF/p75 en IN werd onderzocht om de relatie tussen integratieplaatskeuze, subnucleaire positionering en provirale genexpressie beter te begrijpen. Dit vormt dan ook de rode draad doorheen dit manuscript dat bestaat uit een bundeling van twee verschillende onderzoeksprojecten.

In een eerste deel van mijn werk (zie Hoofdstuk 3), gebruik ik chimere versies van LEDGF/p75 als een tool om integratie te sturen naar potentieel 'veiligere' regio's in het humaan genoom. Hierdoor kunnen ernstige nevenwerkingen geassocieerd met het gebruik van retrovirale vectoren in gentherapeutische toepassingen mogelijks vermeden worden. Ik ontwierp LEDGF/p75 gebaseerde ankers die resulteerden in een meer random integratiepatroon, waarvan algemeen aanvaard wordt dat dit het risico op transcriptionele dysregulatie van naburige genen of 'insertionele mutagenese' beperkt [Chatziandreou et al., 2011; Derse et al., 2007; Kaufmann et al., 2013; Suerth et al., 2012].

Hierna (zie Hoofdstuk 4), bestudeerde ik het belang van integratieplaatsdistributie in de context van HIV-persistentie. In de laatste decennia zorgde de ontwikkeling van antivirale middelen en combinatie AntiRetrovirale Therapie of cART voor een ware revolutie in de behandeling van HIV/AIDS. Het grootste struikelblok ter genezing van HIV blijft echter de aanwezigheid van een stabiel reservoir met transcriptioneel inactief maar replicatiecompetent HIV-provirus dat niet vatbaar is voor cART [Trono et al., 2010]. Deze sluimerende of latente HIV-reservoiren zijn verantwoordelijk voor een heropleving van HIV-viremie reeds enkele weken na de stopzetting van cART. Er wordt algemeen aangenomen dat de integratieplaats en lokale chromatineomgeving een rol spelen in de regulatie van de transcriptionele activiteit van het geïntegreerd provirus. Daarom evalueerde ik of de integratieplaatsselectie mogelijks het reactivatiepotentieel van latent provirus kon beïnvloeden. Meer bepaald ging ik na of LEDGINs, *bona fide* kleine inhibitoren (2-(quinolin-3-yl)azijnzuur-derivaten) die de interactie tussen IN-dimeren en LEDGF/p75 blokkeren en interfereren met de catalytische activiteit van IN [Christ et al., 2010, 2012], het functioneel reservoir kunnen beïnvloeden. Tijdens mijn doctoraat onderzocht ik zowel (i) het tot stand komen van HIV-1-latentie als de (ii) reactivatie van latentie na toediening van 'Latency Reversing Agents' (LRAs).

Deze studies omtrent de rol van de LEDGF/p75-IN wisselwerking in integratieplaatskeuze en HIV-persistentie dragen bij tot de ontwikkeling van veiligere virale vectoren voor gentherapeutische toepassingen ofwel tot een alternatieve strategie voor een HIV-genezing. Hierbij streven we naar een modulatie of preventie van de vorming van het residueel HIV-1-reservoir dat verantwoordelijk is voor de heropleving van HIV-replicatie na het stopzetten van de antivirale behandeling.

Glossary of Terms

A3 or APOBEC3	apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 family
aa	amino acid residues
AAVS1	adeno-associated virus integration site 1
AIDS	acquired immune deficiency syndrome
ART	antiretroviral therapy
ASLV	avian sarcoma leukosis virus
AZT	zidovudine
BET	bromodomain and extra-terminal domain
BMT	bone marrow transplantation
bNAb	broadly neutralizing antibodies
BRD4	bromodomain-containing protein 4
BSA	bovine serum albumin
CA	capsid
CAR	chimeric antigen receptor
cART	combination antiretroviral therapy
CBS	chromatin binding sequence
CCD	catalytic core domain
CCR5	C-C chemokine type 5 receptor
CD4	cluster of differentiation 4
CDK9	cyclin-dependent kinase 9
cDNA	copy deoxyribonucleic acid
CNS	central nervous system
CPSF6	cleavage and polyadenylation specificity factor 6
CR1-4	charged regions 1-4
CRISPR	clustered regularly interspaced short palindromic repeats
CTD	C-terminal domain
CTL	cytotoxic T-lymphocyte
DAPI	4', 6-diamidino-2-phenylindole
DC	dentritic cells
DHS	dnase I-hypersensitive sites
DMEM	dulbecco's modified eagle's medium
DMSO	dimethyl sulfoxide
DTG	dolutegravir

DTT	dithiotreitol
EC*50*	half maximal effective concentration
EDTA	ethylenediaminetetraacetic acid
EFV	efavirenz
eGFP	enhanced green fluorescent protein
ELISA	enzyme-linked immunosorbent assay
EVG	elvitegravir
FACS	fluorescence-activated cell sorting
FDA	U.S. food and drug administration
fLuc	firefly luciferase
GALT	gut-associated lymphatic tissue
GFP	green fluorescent protein
GO	gene ontology
H2A	histone 2A
H2B	histone 2B
H3K4	lysine 4 of histone H3
H3K36	lysine 36 of histone 3
H3K79	lysine 79 of histone 3
H4K16	lysine 16 of histone 4
HDAC	histone deacetylase
HDACi	histone deacetylase inhibitor
HEK293T	human embryonic kidney 293T cells
HEXIM1	hexamethylene bis-acetamide inducible 1
HIV	human immunodeficiency virus
HMT	histone methyltransferase
HP1	heterochromatin protein 1
HPV	human papillomavirus
HRP	horseradish peroxidase
HRP-2	HDGF related protein 2
HSCT	haematopoietic stem cell transplantation
IBD	integrase binding domain
IC*50*	50 % inhibitory concentration
IL	interleukin
IN	integrase
iPS	induced-pluripotent stem cells
IWS1	interacts with Spt6
JPO2	cell division cycle-associated 7-like protein (CDCA7L)
KD	knockdown
KO	knockout

KSHV	kaposi's sarcoma-associated herpes virus
LANA	latency-associated nuclear antigen
LARP7	lupus antigen related protein 7
LB	lysogeny broth
LEDGF	lens epithelium derived growth factor
LRA	latency reversing agents
LTR	long terminal repeat
LV	lentiviral vector
MA	matrix
MFI	mean fluorescence intensity
MLL	myeloid/lymphoid leukemia or mixed-lineage leukemia
MLV	murine leukemia virus
MMTV	mouse mammary tumour virus
MRC	matched random control
mRNA	messenger RNA
NC	nucleocapsid
NCINI	non-catalytic site integrase inhibitor
Nef	negative regulatory factor
NF$\varkappa$B	nuclear factor $\varkappa$B
NLS	nuclear localization signal
NNRTI	non-nucleoside reverse transcriptase inhibitor
NRTI	nucleoside/nucleotide reverse transcriptase inhibitor
NTD	N-terminal domain
NUP	nucleopore
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEI	polyethylenimine
PFV	prototype foamy virus
PI	protease inhibitor
PIC	pre-integration complex
PogZ	pogo transposable element with zinc finger domain
PrEP	pre-exposure prophylaxis
PSIP1	PC4- and SFRS1-interacting protein 1
P-TEFb	positive transcription elongation factor b
PV	papillomavirus
PWWP	Pro-Trp-Trp-Pro domain

qPCR	quantitative PCR
RNAi	RNA interference
RPMI	roswell park memorial institute medium
RT	reverse transcriptase
RT qPCR	real time quantitative PCR
RV	retroviral vector
SDS	sodium dodecyl sulfate
SFFV	spleen focus-forming virus
SI	selectivity index
SIN	self-inactivating
SIV	simian immunodeficiency virus
SRD	supercoiled recognition domain
START	strategic timing of antiretroviral treatment
SU	surface unit
TALEN	transcription activator-like effector nucleases
TAR	transactivation response element
Tat	transactivator of transcription
TM	transmembrane unit
TSS	transcription start site
UCR	ultra conserved regions
vDNA	viral deoxyribonucleic acid
Vif	virus infectivity factor
VOA	viral outgrowth assay
Vpr	viral protein R
Vpu	viral protein unique
vRNA	viral ribonucleic acid
VSV-G	vesicular stomatitus virus glycoprotein
WAS	wikott-aldrich syndrome
WT	wild-type
WPRE	woodchuck hepatitis virus post-transcriptional regulatory element
X-SCID	X-linked severe combined immunodeficiency
X-CGD	X-linked chronic granulomatous disease

Table of Contents

List of Figures

List of Tables

1 General Introduction

For more than 10 years our research group has been focusing on cellular cofactors that support Human Immunodeficiency Virus (HIV) replication. HIV is the prototype of the *Lentiviridae* which are characterized by their ability to stably integrate a viral DNA copy into the host cell genome and efficiently infect both dividing and non-dividing cells. These characteristics qualify lentiviral vectors as attractive gene transfer vehicles for a range of gene therapeutic applications. In 2003, the Lens Epithelium-Derived Growth Factor/p75 (LEDGF/p75) was identified as the major cellular cofactor orchestrating lentiviral integration and integration site selection. This manuscript encompasses two lines of research with a main focus on LEDGF/p75. In a first part I modified LEDGF/p75 to redirect lentiviral integration, improving the lentiviral vector genotoxicity profile for gene therapeutic purposes (see Chapter 3). In a second part, I studied the effect of the LEDGF/p75-IN interplay, lentiviral integration site distribution and nuclear topology on HIV latency and reactivation (see Chapter 4). Both studies are preceded by a general introduction on HIV molecular biology.

1.1 HIV, the causative agent of AIDS

1.1.1 HIV epidemiology

Human Immunodeficiency Virus (HIV) is a human pathogen and the causative agent of the Acquired Immune Deficiency Syndrome or AIDS. A series of reports on clustered cases of *Pneumocystis carinii* pneumonia [Centers for Disease Control (CDC), 1981] and Kaposi's sarcoma [Hymes et al., 1981] among homosexual men in the early 80's triggered global awareness for a till then unknown disease. Few more years after these reports the research groups of Prof. Luc Montagnier [Barre-Sinoussi et al., 1983] (Institute Pasteur, Paris) and Prof. Robert Gallo [Gallo et al., 1983] (National Cancer Institute, Washington D.C.) isolated a new virus strain, coined Lymphadenopathy-Associated Virus or Human T-Lymphotropic Virus III respectively, acting as the responsible etiological agent of AIDS. Both viruses were shown to be identical and later renamed Human Immunodeficiency Virus (HIV) in 1986. The HIV virus originated from separate zoonotic transmissions with SIV (Simian Immunodeficiency Virus) strains from African primate species. HIV type-1 was transmitted from chimpanzee [Gao et al., 1999], whereas HIV type-2 originated from sooty mangabey monkeys [Sharp et al., 2001]. The HIV type-1 accounts for the current global HIV pandemic and almost invariably results in immunodeficiency while HIV-2 is less transmissible and infection progresses more slowly towards AIDS. Since the start of the AIDS pandemic in the early 1980s and until recently, annual AIDS deaths have been steadily increasing. Today over 36 million people across the world are living with HIV (UNAIDS, 2015) of which the majority are living in Sub-Saharan Africa. 2.1 million new infections were estimated in 2015 with over 1 million annual deaths from AIDS-related causes. HIV-1 infection in humans is considered to be pandemic by the World Health Organization (WHO, Global summary of the AIDS epidemic, 2010)). Introduction of effective combination AntiRetroviral Therapy (cART) in the late 1990's is able to suppress plasma viral loads below the level of detection and dramatically reduced mortality and morbidity associated with HIV-1 infection. Yet currently, merely half of the diagnosed cases worldwide have access to antiretroviral

therapy (17 million). However, despite the success of cART, persistence of latent, replication-competent HIV provirus remains the major impediment towards a cure of HIV/AIDS.

1.1.2 Retroviral taxonomy

HIV is an enveloped RNA virus enclosing a diploid, single-stranded, positive RNA genome and belongs to the family of the *retroviridae* (Retroviruses). A hallmark of these viruses is their ability to reverse transcribe their RNA genome into a double stranded DNA copy which gets stably incorporated into the host cell genome by the viral integrase enzyme (IN). This integration event irreversibly links the fate of the invading virus to that of its host cell. Depending on their genetic complexity retroviruses are generally categorized as being simple or complex. The subfamily of the *Orthoretrovirinae* to which HIV belongs can be divided into six distinct genera based on the phylogenetic alliance of the highly conserved *pol* gene: the *Alpha-*, *Beta-*, *Gamma-*, *Delta-*, *Epslilonretroviridae* and *Lentiviridae*. The subfamily of *Spumaretrovirinae* only consists of the *Spumaviridae*. Classification based on *pol* similarity is congruent with a number of other structural and biological traits; significant differences are found in genetic complexity, host range and host interaction. HIV belongs to the *Lentiviridae*, with *lentus* meaning slow referring to the long incubation period between initial infection by an HIV founder virus and the onset of clinical symptoms of AIDS (see Section 1.1.3). These lentiviruses harbor the unique capability to infect non-dividing cells in addition to dividing cells. Different HIV-1 strains can be further classified into genetically diverse subgroups M, N, O and P, of which M is the most prevalent.

1.1.3 HIV pathophysiology

In general, HIV transmission occurs via semen or vaginal fluid upon sexual contact, via infected blood (-products) or in a vertical manner from mother to child. Disease progression towards AIDS can be roughly divided in three phenotypically distinct phases: I) acute seroconversion, II) asymptomatic infection (or clinically latent phase) and III) AIDS. Acute infection (phase I, Figure 1.1 panel a on page 4) is initiated at the mucosal membrane where a first infected focal founder population of $CD4^+$ T-cells allows for rapid viral expansion and a boost in plasma viral loads causing a strong inflammatory response (Figure 1.1 panel b). This acute phase can be associated with flu-like symptoms and lymphadenopathy. Of note, early during infection a reservoir of latently infected cells is established mainly residing in long-lived memory $CD4^+$ T-cells. The subsequent immune system clearance (innate and adaptive immune responses) contributes to the establishment of a steady state level (asymptomatic phase II, Figure 1.1 panel a). Here, infected individuals show relatively few or no symptoms even though ongoing viral replication occurs. This phase is typically characterized by a slow (*lentus*) decline in $CD4^+$ T-cells (the predominant target cells of HIV) and disease progression. The decrease in $CD4^+$ T-cells is a consequence of both immune hyperactivation driving T-cell apoptosis and viral cytophatic effects [Goonetilleke et al., 2009]. Over time mutations in key epitopes emerge often leading to immune escape as a consequence of the error prone HIV RT function. The final phase III is characterized by AIDS (Figure 1.1 panel a) usually taking place around 10 years post infection and is diagnosed when $CD4^+$ T-cell counts shrink

to less than 200 cells/µl. Deregulation of the $CD4^+/CD8^+$ T-cell balance leads to an improper immune response, unable to adequately counteract opportunistic infections such as e.g. Kaposi's sarcoma or Tuberculosis (TB) which remains the leading cause of death among people living with HIV.

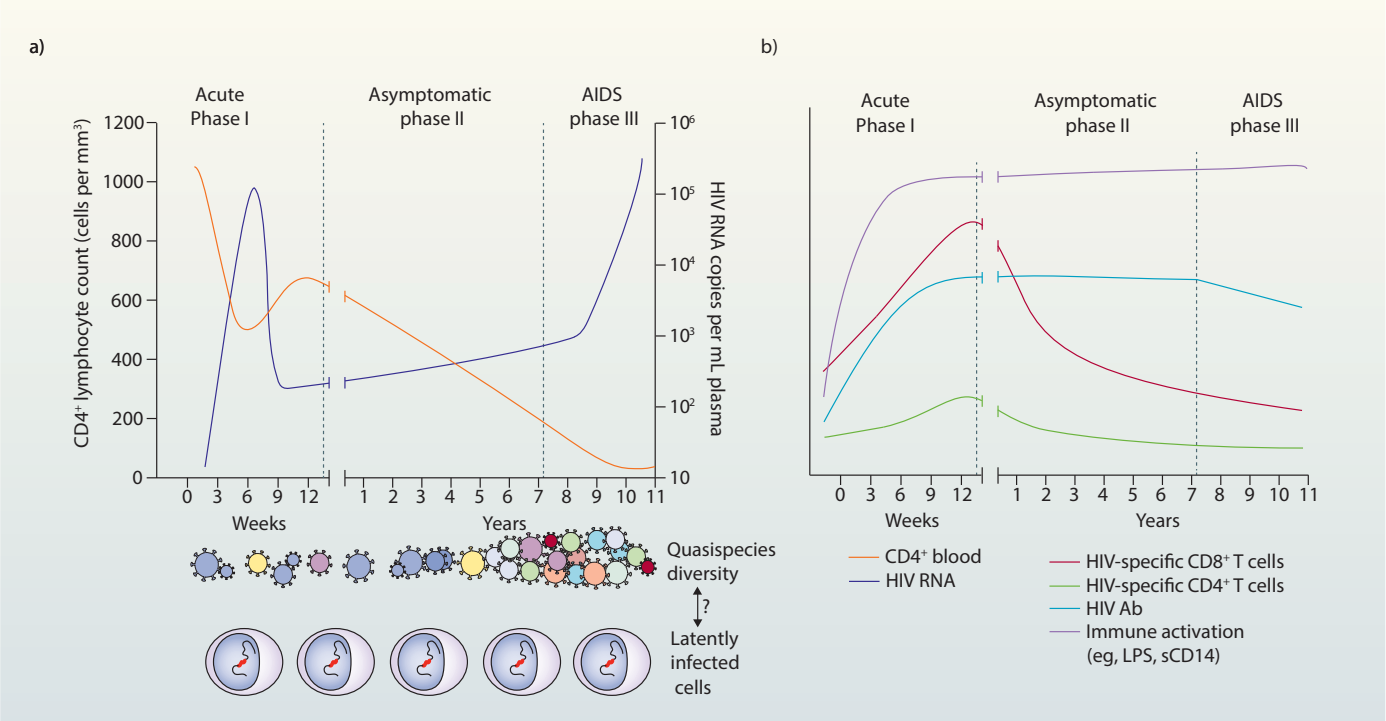


Figure 1.1: Disease progression from viral infection to the development of AIDS. Panel a) depicts the $CD4^+$ T-cell counts and plasma viral load changes over time post HIV founder infection for untreated patients. During the course of infection viral escape mutants arise allowing HIV to bypass the immune surveillance. The reservoir of latently infected cells is established already early during infection and persists during the life time of the patient (mainly in resting memory $CD4^+$ T-cells). Panel b) depicts the acute immune response to HIV infection as represented by the changes in immune activation and production of non-neutralizing antibodies and HIV specific T-cells. CD, cluster of differentiation; LPS, lipopolysaccharide. *(Modified from [Maartens et al., 2014])*

1.1.4 HIV particle structure and genomic organization

An enveloped HIV virion has a spherical morphology with a diameter of roughly 100 to 120 nm (depicted in Figure 1.2 a on page 5). The phospholipid envelope bilayer contains the viral surface glycoprotein gp120 (SU) and transmembrane protein gp41 (TM) forming the viral spikes or cap (composed of three gp120 units and three gp41 units). The inner surface of the envelope is lined by the matrix (MA or p17) protein which encloses a cone-shaped core, constituted of hexamers and pentamers of CA protein or p24 [Pornillos et al., 2011]. The viral core encompasses two positive single strands of a 9.7 kb RNA genome along with viral (IN, RT, PR, Nef, Vpr and Vif) and cellular proteins. This viral diploid RNA genome is covered by nucleocapsid (NC or p7) proteins. The HIV-1 genome (Figure 1.2 b, [Foley et al., 2015]) is flanked on both sides by long terminal repeats (LTRs) initially consisting of a 5' unique element (U5) and a repeat element (R) at the 5' end and a 3' unique element (U3) and R at the 3' end. Following

reverse transcription both LTRs contain U3-R-U5, the full-length LTR. Like all other *Retroviridae*, HIV-1 contains three major Open Reading Frames (ORFs); *gag*, *pol* and *env*. These ORFs encode for polyproteins that are subsequently cleaved into mature proteins. *Gag* encodes for the structural proteins MA, CA and NC while *pol* encodes for the viral enzymes PR, RT and IN. *Env* encodes for the gp120 and gp41 glycoproteins. In addition, complex retroviruses like lentiviruses encode for several accessory proteins. In the case of HIV-1 six accessory proteins are present: Vif, Vpr, Tat, Rev, Vpu and Nef (reviewed in [Simon et al., 2015]). Vif or virion infectivity factor is able to counteract the cellular restriction factor APOBEC3 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3), a single-stranded cytidine deaminase able to induce GG to AG or GA to AA hypermutations. Viral protein R (Vpr) affects cell cycle arrest and apoptosis. Tat or transactivating protein is the major determinant of the positive feedback circuitry driving viral transcription. Regulator of viral protein expression (Rev) binds to the Rev responsive element (RRE), stabilizing unspliced and singly spliced viral RNA and enhancing nuclear export to the cytoplasm. Vpu or unique viral protein stimulates CD4 receptor degradation. Both Vpu and Nef (negative regulatory factor) prevent tetherin (BST2) from holding on to the budding virus at the plasma membrane.

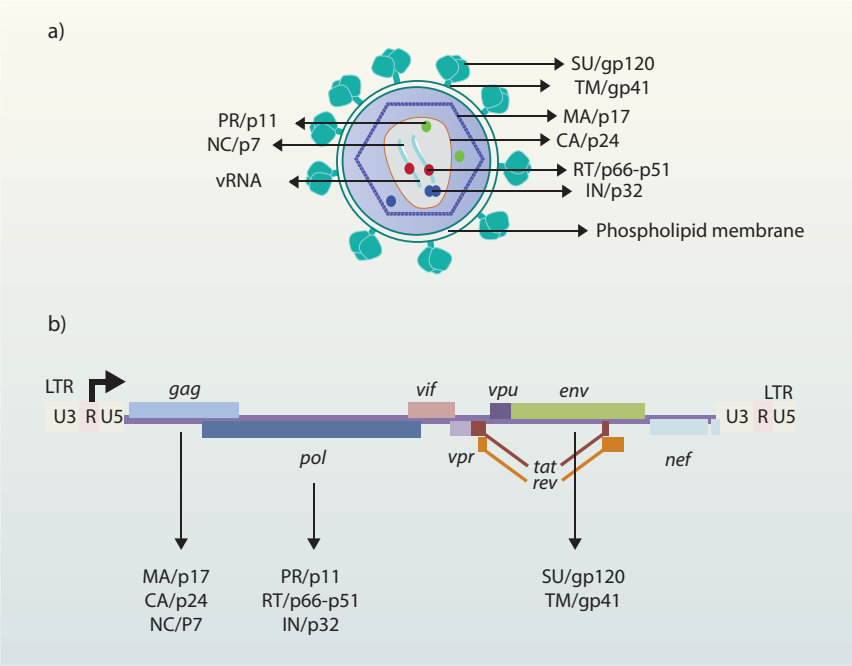


Figure 1.2: HIV-1 particle structure and genomic organization. a) Diagram of an HIV virion depicting both structural and enzymatic constituents. b) Genetic organization of HIV, a complex retrovirus. The HIV-1 proviral genome is flanked by long terminal repeats (LTRs) at both sides. Like all other *Retroviridae*, HIV-1 contains three major open reading frames; *gag*, *pol* and *env*. *Gag* encodes for structural proteins MA, CA and NC. *Pol* encodes for the viral enzymes PR, RT and IN. *Env* encodes for the gp120 and gp41 glycoproteins. Complex retroviruses like lentiviruses additionally encode for accessory proteins. In the case of HIV-1 these accessory proteins are the virion infectivity factor (Vif), the viral protein R (Vpr), the transactivating protein (Tat), Regulator of viral protein expression (Rev), unique viral protein (Vpu) and negative regulatory factor (Nef). PR, protease; NC, nucleocapsid; SU, surface envelope glycoprotein; TM, transmembrane protein; MA, matrix; CA, capsid; RT, reverse transcriptase; IN, integrase; LTR, long terminal repeat.

1.1.5 HIV replication cycle

In order to replicate in their hosts, viruses need to navigate around the complexities of the cell, co-opting mechanisms of cellular physiology while defeating restriction factors able to halt their proliferation. The success of HIV-1 stands in contrast to its apparent simplicity. Even though its small 9.7 kb genome only encodes 15 mature proteins (depicted in Figure 1.2 on page 5) the virus is able to elegantly employ the infected cell and subvert the innate and adaptive immune responses of the host, resulting in a persistent infection in humans. Viral replication proceeds through a number of characteristic steps, depicted in Figure 1.3 on page 7. Briefly, HIV predominantly infects $CD4^+$ T-cells like T-lymphocytes, macrophages and dendritic cells through initial attachment (1) of the gp120 spike proteins to the cellular CD4 receptor. The latter allows interaction with the C-X-C chemokine receptor type 4 (CXCR-4, R4 tropic HIV) or C-C chemokine receptor type 5 (CCR5, R5 tropic HIV) co-receptors which induces a conformational change that triggers membrane fusion (2) of the viral particle with the host cell membrane. Of note, for research purposes HIV particles are often pseudotyped with a Vesicular Stomatitis Virus Glycoprotein (VSV-G) envelope which broadens the cell tropism but leads to a different entry mode. Upon fusion the viral core is released into the cell cytoplasm. *En route* towards the nucleus the virus tightly interacts with the cellular cytoskeleton and moves across the filament network [McDonald et al., 2002]. During this translocation the capsid core uncoats (3) and a double stranded DNA copy of the viral RNA genome is created by the reverse transcriptase (RT, 4). A PreIntegration nucleoprotein Complex (PIC) is formed by association of both viral and cellular proteins and is actively transported into the nucleus via the Nuclear Pore Complex (NPC) (5). Subsequently, the vDNA intermediate will become integrated in the host chromosomal DNA by means of the viral integrase enzyme, forever linking the fate of the invading provirus to that of its host cell (6). The integrated provirus will adopt a silent/quiescent state or actively produce new virions. HIV therefore hijacks the normal transcription and translation machinery of the host cell. Early multiply spliced RNA transcripts encode for accessory proteins Tat and Rev. A Tat regulated positive feedback loop stimulates RNA Pol II mediated HIV transcription by binding to the TransActivation Response element (TAR) in the viral LTR promoter and attracts the Positive Transcription Elongation Factor-b (P-TEFb). Rev mediates nuclear export of both unspliced and singly spliced RNA leading to the production of Gag, Pol, Env. These polyproteins align then at the cellular membrane and assemble into budding viral particles (10, 11). Upon release the viral particles will undergo proteolytic maturation induced by the viral PRotease enzyme (PR), producing novel infectious virions (12, 13).

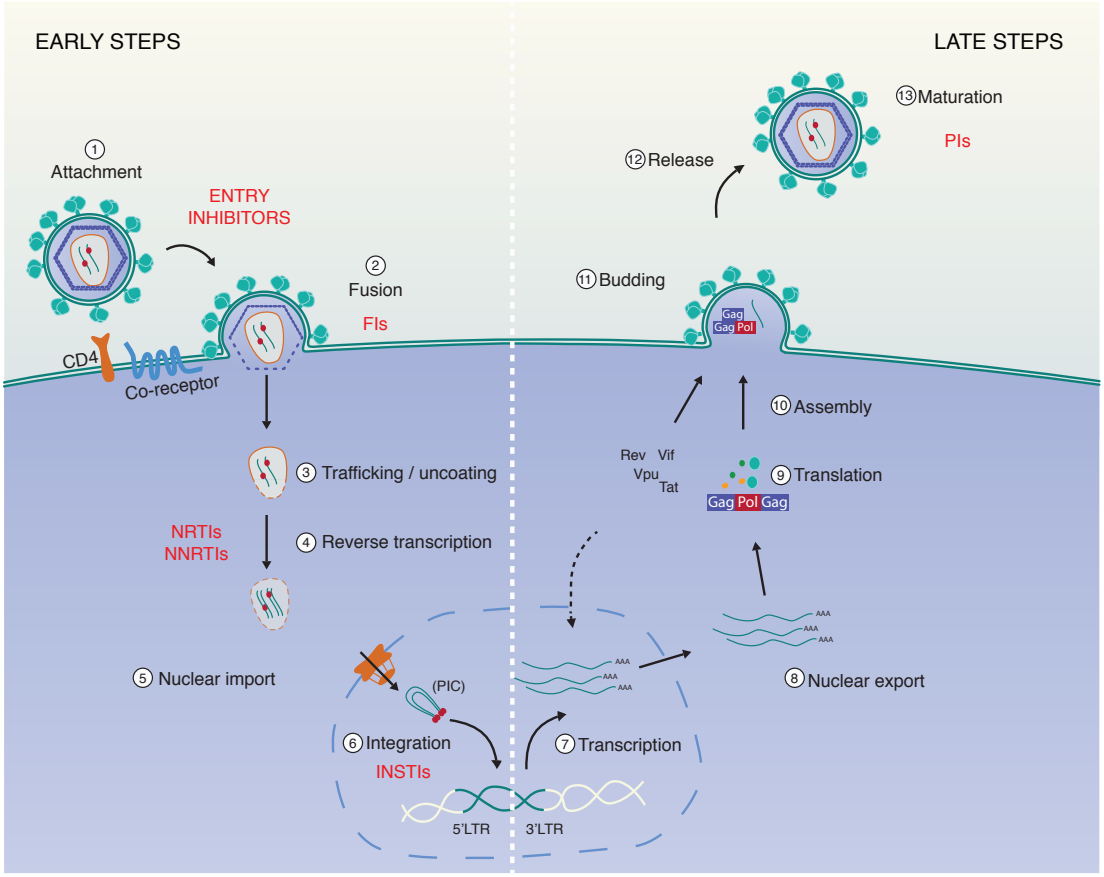


Figure 1.3: HIV replication cycle. Schematic representation of both 'early' and 'late' steps in the HIV replication cycle together with the antiretroviral drug classes acting on the respective processes. Upon Env-mediated attachment of the virion to the host cell CD4 receptor and CXCR4 or CCR5 co-receptor, a fusion of the viral envelope with the cell membrane occurs. Fusion triggers a release of the conical capsid core into the cytoplasm. On its way to the nucleus the capsid disassembles and RT creates a double-stranded DNA copy of the viral genome. Various viral and endogenous proteins will assemble into a PreIntegration Complex (PIC) which is actively imported into the nucleus. The viral integrase catalyzes the insertion of the vDNA into the host chromatin. Proviral transcription, export and translation of viral mRNA leads to the production of novel virions. These particles are released from the cell and mature via polyprotein processing by the viral protease. FIs, fusion inhibitors; NRTIs, nucleoside/nucleotide reverse transcriptase inhibitors; NNRTIs, non-nucleoside reverse transcriptase inhibitors; INSTIs, integrase strand transfer inhibitors; PIs, protease inhibitors.

1.1.6 Retroviral integration

Upon nuclear entry the PIC encounters the host cell chromatin. The viral integrase inserts the reverse transcribed viral genome, a point of no return (illustrated schematically in Figure 1.4 on page 8 and reviewed in [Li et al., 2011]). Hereby the fate of the invading virus is forever linked to that of its host cell, persisting for the lifetime of the infected cell or its progeny. Briefly, a dimer of IN enzymes bound to each viral LTR end is required to remove a dinucleotide (GT in the case of HIV) from both 3' LTR ends (upon synthesis of the ds vDNA), a process referred to as 3' processing (3P, Figure 1.4 a and b). Removal of the terminal GT nucleotides exposes a nucleophilic 3'-hydroxyl group on the conserved CA dinucleotide. The retroviral intasome will next capture the host target DNA in a specific manner, showing a weak palindromic sequence preference owing to the intasome central dyad axis (Figure 1.4

c and subsubsection 1.1.6.2). The processed vDNA with the recessed LTR ends, now in complex with a multimer of IN proteins (intasome, at least a tetramer [Hare et al., 2010]), will be inserted in the host cell genome via a transesterification reaction (Figure 1.4 d). This Strand Transfer reaction (ST) results in a gapped intermediate where both vDNA strands are joined by their 3' ends to the opposing strands. The insertion sites are separated by 4-6 bp (reviewed in [Engelman and Cherepanov, 2014]). These corresponding gaps are repaired by the host cell DNA repair machinery (Figure 1.4 e) leading to provirus establishment.

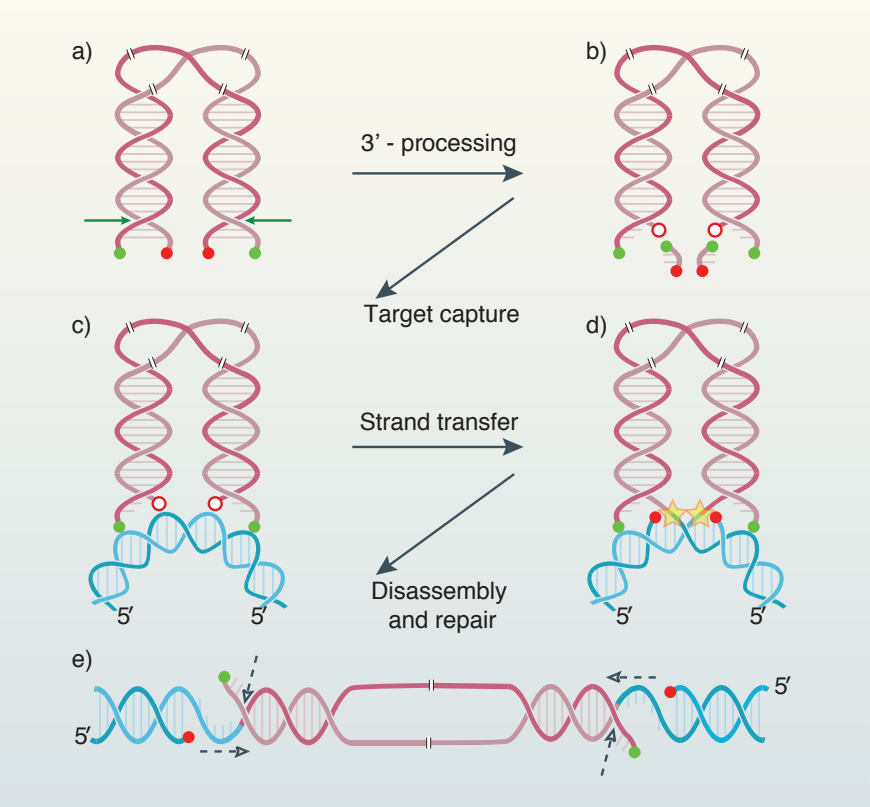


Figure 1.4: Schematic overview of the lentiviral integration process. (a) 3' processing takes place on the way to the nucleus following reverse transcription. (b) The water-mediated endonucleolytic cleavage of the viral DNA (green arrow), 3' from the conserved CA dinucleotides, generates a reactive 3'-hydroxyl group at both viral ends (indicated with red circles); The IN - vDNA assembly, known as the intasome, captures the target DNA (blue) (c) and carries out the strand transfer reaction (yellow stars (d), 3'-end joining). The two vDNA 3'-hydroxyl ends (red circles) attack a DNA phosphodiester bond on each strand of the host DNA acceptor with a five-base-pair stagger across the DNA major groove. (e) The removal of unpaired 5'-ends of the viral DNA (arrows in e) and gap filling are done by cellular repair enzymes. All other 3'-hydroxyl ends and 5'-phosphate ends are shown as red and green dots, respectively. *(Modified from [Pommier et al., 2005])*

1.1.6.1 Retroviral integrase and intasome structure

The lentiviral IN is a 32 kDa enzyme encoded by the Pol ORF. All retroviral IN proteins share three structurally conserved domains which are interconnected by flexible linker regions: (1) the amino-Terminal Domain (NTD) (2) the Catalytic Core Domain (CCD) and (3) the Carboxy-Terminal Domain (CTD) (Figure 1.5 on page 9, reviewed in [Jaskolski et al., 2009; Cherepanov et al., 2011]). Together

these domains are involved in DNA capture and IN multimerization and required for its biological activity. The NTD contains a His_2-Cys_2 (HH-CC) Zn^{2+} coordinating motif involved in stabilization of the triple helix bundled structure (Figure 1.5 left panel, [Zheng et al., 1996; Cai et al., 1997; Eijkelenboom et al., 2000]). The CCD is highly conserved among different *Retroviridae* and contains the catalytic DDE triad, which is responsible for the coordination of two Mg^{2+} or Mn^{2+} ions necessary for its activity (Figure 1.5 middle panel). The CCD domain is composed of five β-sheets surrounded by several α-helices. The CTD forms a 5-stranded β-barrel structure (Figure 1.5 right panel).

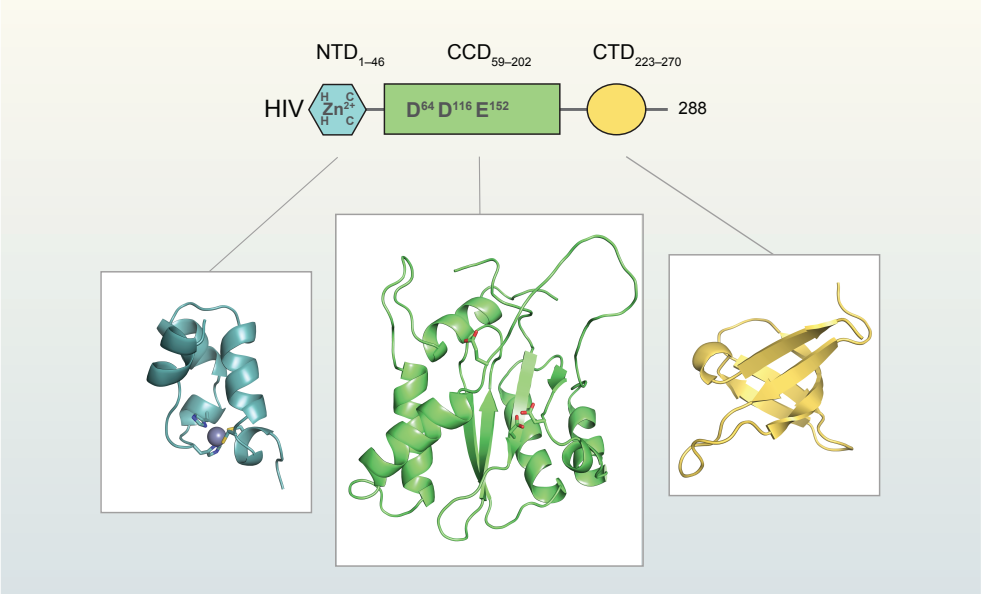


Figure 1.5: Functional domains of the lentiviral integrase. Depiction of the different HIV-1 integrase domains; Left panel depicts the tertiary structure of the N-Terminal Domain, middle panel the Catalytic Core Domain and right panel the C-Terminal Domain (NTD, CCD and CTD, respectively) *(Modified from [Demeulemeester et al., 2015])*

Unfortunately, full length HIV IN has been refractory to structural studies. However, structural homology among retroviral IN orthologues has allowed their use as a proxy for HIV IN [Hare et al., 2010]. The crystal structure of the prototype foamy virus IN (PFV, *Spumaviridae*, Figure 1.6 a & b on page 10) therefore represented a landmark in the field of retrovirology and unveiled some of the workings of the integration machinery. Here, the intasome complex consists of a dimer of IN dimers [Hare et al., 2010; Maertens et al., 2010; Hare et al., 2012; Maskell et al., 2015]. The two inner subunits are responsible for bridging both dimers and are making extensive contacts with the vDNA [Hare et al., 2010]. Only the inner dimer is responsible for the integration reaction while the outer IN dimers interact solely through the CCD. The target DNA (tDNA) is accommodated in the cleft between the active sites, displaying a severely bent conformation. Positioning of the active sites allows targeting of the phosphodiester bonds in the major groove [Maertens et al., 2010]. This structural insight explains the preference of IN for host DNA sites that allow a minimal bendability such as nucleosomal DNA [Pryciak and Varmus, 1992; Pruss et al., 1994b,a; Maskell et al., 2015]. More recently, three-domain crystal structures of the Rous Sarcoma Virus (RSV, *Alpharetroviridae*, [Yin et al., 2016]) and Mouse Mammary Tumour Virus (MMTV, *Betaretroviridae*, Figure 1.6 c [Ballandras-Colas et al., 2016]) together with vDNA indicated

a novel octameric intasome assembly and highlights a remarkable diversity among retroviruses. Here, the core is formed by a pair of proximal IN dimers interacting with the vDNA ends and engagement of two flanking non-catalytic IN dimers via their CTD domain bridging the interaction in *trans*. Their respective composition differs in linker length between the CCD and the CTD. For PFV this linker is comprised of 50 aa residues, which affords positioning of the the inner CTD for both viral and target DNA binding. In MMTV and RSV this linker is only a mere 8 aa. Therefore, additional flanking IN dimers are required positioning their CTD in the intasome core. HIV-1 IN however, contains an intermediate linker length of 20 aa, fueling the ongoing debate about the exact IN stoichiometry in the HIV intasome complex responsible for concerted integration.

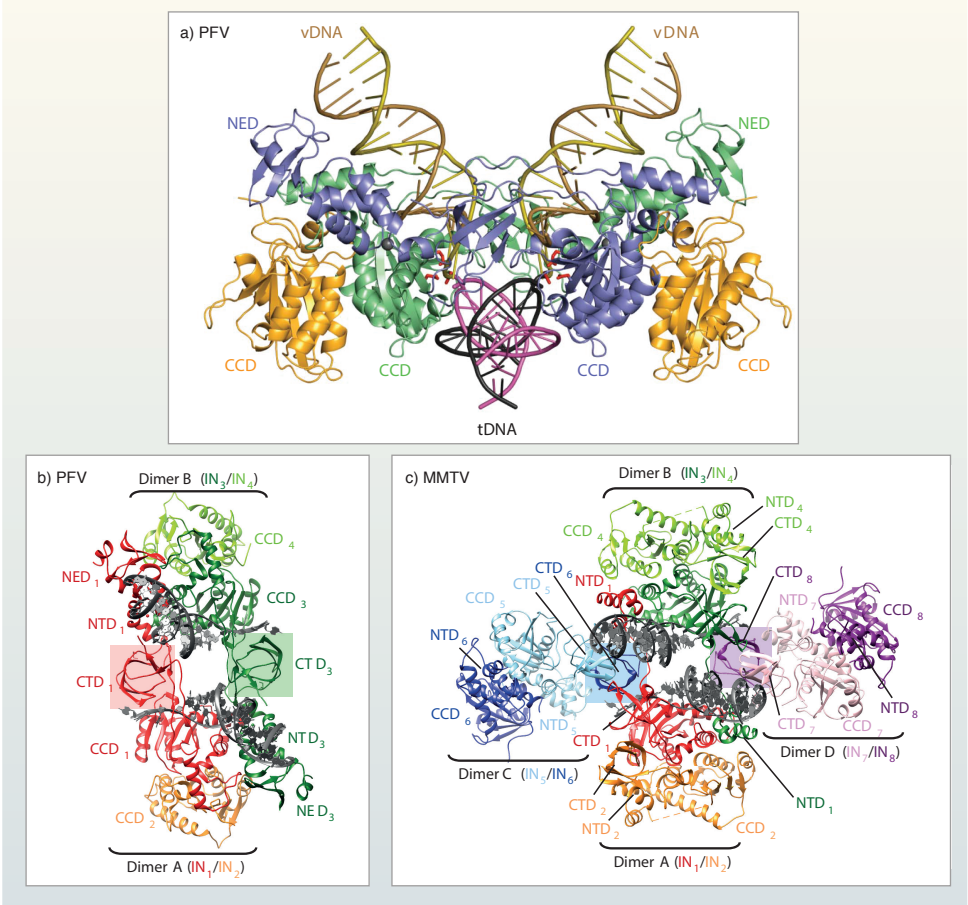


Figure 1.6: Retroviral intasome structures. a) Perpendicular representation of the PFV intasome in complex with the viral and target DNA. Inner subunits form the active cleft (green and blue). b and c) PFV and MMTV intasome architecture respectively. Core dimers A and B are colored red/orange and green/lightgreen while for MMTV the flanking IN dimers C and D are colored blue/skye blue and purple/light pink. NTD, N-Terminal Domain; CCD, Catalytic Core Domain; CTD, C-Terminal Domain; NED, N-terminal Extension Domain. *(Modified from [Maertens et al., 2010] and [Ballandras-Colas et al., 2016])*

1.1.6.2 Determinants of retroviral integration site selection

At first, it was accepted that retroviruses integrated close to random throughout the genome of the invaded cells. However, during the last decade, it became clear that retroviral integration is a non-random, multi-step process in which biases are introduced at different levels leading to an integration profile distinct for all retroviral genera, each favoring specific chromatin environments (reviewed in [Demeulemeester et al., 2015]). Lentiviruses, such as the Human Immunodeficiency Virus Type-1 (HIV-1) for instance preferentially integrate into the body of actively transcribed regions [Schroder et al., 2002]. Conversely, gammaretroviruses such as the Murine Leukaemia Virus (MLV) target active promoter-proximal as well as distal enhancer elements [Wu et al., 2003; Mitchell et al., 2004; LaFave et al., 2014]. Avian Sarcoma-Leukosis Virus (ASLV, Alpharetroviruses) maintains a preference for open chromatin regions, but shows no bias for transcriptional regulatory elements or transcription units [Moiani et al., 2014]. A first asymmetry is introduced by the nuclear entry route (Figure 1.7 panel a on page 12). Lentiviruses evolved the unique ability to traverse nuclear pore complexes (NPC) allowing them to infect non-dividing cells while other viruses, such as MLV, evolved mechanisms to segregate along with the chromosomes during mitosis [Elis et al., 2012]. The lentiviral capsid core is believed to dock onto the cytoplasmic side of the NPC through interactions with Nup358-RanBP2. Engagement of other nucleoporins such as Nup153 and Nup98-Nup96, additional import factors such as Transportin-SR2 (TRN-SR2, TNPO3) or cellular cofactors such as Cleavage and Polyadenylation Specificity Factor 6 (CPSF6) orchestrate the complex process of size limited import of the PIC to the nucleoplasm. This implies that the first nuclear subcompartment a lentiviral PIC encounters is the NPC-associated cone-like Heterochromatin Exclusion Zones (HEZs) near their nuclear baskets [Pascual-Garcia and Capelson, 2014; Sood and Brickner, 2014] adjacent to the heterochromatin amassment of Lamin-Associated Domains (LADs, [Guelen et al., 2008]). Different fluorescent microscopy studies reported that the HIV PIC as well as the provirus are located in euchromatin regions in proximity of the nuclear rim [Albanese et al., 2008; Di Primio et al., 2013] as opposed to MLV, which distributes more randomly throughout the nucleus [Quercioli et al., 2016].

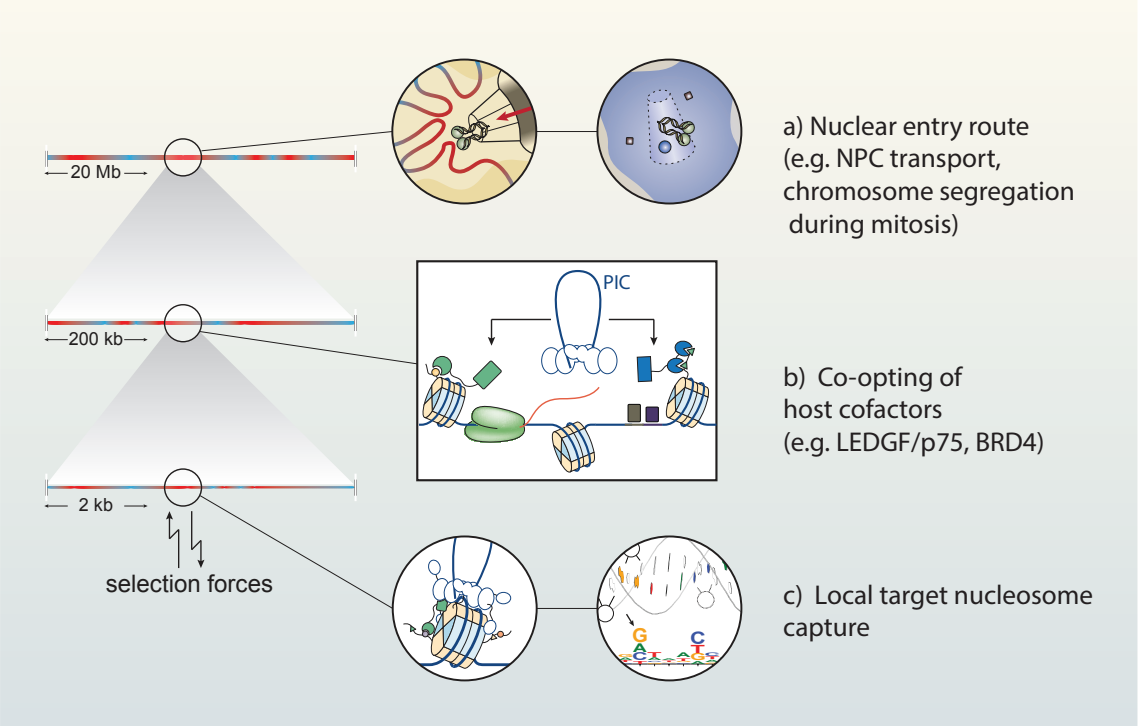


Figure 1.7: Overview of retroviral integration targeting parameters. Retroviral integration is a non-random, multi-step process in which biases are introduced at different levels. a) Taking a specific route into the nucleus (nuclear import or chromosome segregation upon mitosis) generates a first asymmetry. b) Hijacking of cellular cofactors further defines tethering towards specific chromatin regions. c) Capture of the nucleosomal target DNA by the intasome, catalyzing the integration process, results in an additional contribution to the overall integration site selection. NPC, Nuclear Pore Complex; LEDGF/p75, Lens Epithelium-Derived Growth Factor p75; BRD4, Bromodomain Containing protein 4. *(Modified from [Demeulemeester et al., 2015])*

Next the PIC co-opts cellular cofactors tethering it to particular chromatin contexts (Figure 1.7 panel b). In the case of lentiviruses, Lens Epithelium Derived Growth Factor/p75 (LEDGF/p75) is responsible for directing integration towards the body of active transcription units [Schroder et al., 2002; Maertens et al., 2003] by direct interaction with the viral IN and specific histone tails (H3K36me3). Conversely, gamma-retroviral PICs adopt Bromodomain and ExtraTerminal domain (BET) proteins targeting integration towards active promoter-proximal as well as distal enhancer elements after chromosome segregation [De Rijck et al., 2013; Gupta et al., 2013; Sharma et al., 2013].
A third parameter determining integration site choice is the molecular recognition between the retroviral intasome and the local target nucleosome (Figure 1.7 panel c). Recent studies employed the PFV intasome structure to predict and modify IN-tDNA contacts in the context of HIV-1 infection [Serrao et al., 2014; Demeulemeester et al., 2014b]. Two IN-aa form close contacts with the tDNA base (HIV-1 IN_{S119} and IN_{R231}) and directly influence the local palindromic sequence preference owing to the central dyad axis. Different amino acids at these two positions yield distinct local sequence biases for viral integration and modulate central tDNA bending. Surprisingly, HIV-1 IN_{S119G} and IN_{R231G} mutations at these positions also shifted global integration preference away from gene-dense regions which could possibly be explained by different tDNA bending requirements which consequently reduces the shape and/or electrostatic compatibility with certain nucleosomes.

1.1.6.3 LEDGF/p75, an HIV IN cellular cofactor

Lens Epithelium Derived Growth Factor p75 or LEDGF/p75 is a ubiquitously expressed chromatin reader of 530 aa encoded by the PC4- and SFRS1-interacting protein 1 gene (*PSIP1*) together with its smaller splice variant LEDGF/p52 (333 aa), which is located on chromosome 9p22.2 (Figure 1.8 on page 14, [Ge et al., 1998]). LEDGF/p75 has been identified as the main cellular cofactor directing lentiviral integration towards the body of actively transcribed genes and was initially discovered as a binding partner of the lentiviral IN in 2003 [Cherepanov et al., 2003]. Meanwhile LEDGF/p75 has been thoroughly validated as an essential HIV cofactor (for a review see [Debyser et al., 2015]).
LEDGF/p75 is a bimodal protein and a member of the Hepatoma-Derived Growth Factor (HDGF) family, characterized by its N-terminal PWWP domain corresponding to the Pro-Trp-Trp-Pro motif (aa 1-91 Figure 1.8, [Izumoto et al., 1997; Stec et al., 2000]). This domain recognizes nucleosomal DNA and specific methyl lysines on histone tails (H3K36me3), present in transcriptionally active chromatin regions [Eidahl et al., 2013; van Nuland et al., 2013]. In addition to the chromatin reading domain, LEDGF/p75 contains several additional DNA binding motifs; a nuclear localization signal (NLS, aa 148-156 [Maertens et al., 2004]), two A/T hook-like elements (aa 178-199) [Maertens et al., 2004; Llano et al., 2006] and 4 charged regions (CR1-4) overlapping with a negatively super-coiled DNA recognition domain (aa 206-336, CR-2-4) [Tsutsui et al., 2011]. Aside from these N-terminal chromatin/DNA interacting domains, LEDGF/p75 also harbors a unique C-terminal protein binding domain (PBD, aa 347-454), which is absent in LEDGF/p52. In contrast, LEDGF/p52 contains a small C-terminal tail of 8 aa residues. The PBD is constituted of a compact right-handed bundle of 5 α-helices which has been coined Integrase Binding Domain or IBD [Cherepanov et al., 2004]. The main interaction interface with the lentiviral IN is formed between the IBD of LEDGF/p75 and the catalytic core domain of IN dimers. Binding of LEDGF/p75 to the viral integrase dimer interface modulates IN multimerization and catalytic activity [Cherepanov et al., 2003], protecting IN from proteolytic degradation [Llano et al., 2004]. Next to LEDGF/p75 the paralogue Hepatoma Derived Growth Factor-Related Protein 2 (HRP-2, *HDGFRP2*) is the only additional endogenous protein carrying both such PWWP and IBD domains (Figure 1.8, [Cherepanov et al., 2004]). HIV IN hijacks these host cell proteins in order to tether the viral Pre-Integration Complex (PIC) towards the host cell chromatin [Schrijvers et al., 2012b; Wang et al., 2012]. Depletion of LEDGF/p75 reduces HIV infection and shifts integration out of transcription units, a phenotype which is more pronounced in LEDGF/p75 KO cells [Ciuffi et al., 2005; Shun et al., 2007; Marshall et al., 2007; Schrijvers et al., 2012b], although still not completely random. This bias in part can be ascribed to HRP-2. Only in the abscence of LEDGFp/75, HRP-2 orchestrates integration site selection [Schrijvers et al., 2012b; Wang et al., 2012]. Simultaneous knock down of both proteins further reduces, but not completely abolishes, integration in transcription units [Schrijvers et al., 2012b,a]. Biophysical studies revealed a scan-and-lock mechanism for chromatin tethering of LEDGF/p75 [Hendrix et al., 2011]. The remaining preferences are determined by the intrinsic IN properties and for a major part by nuclear topology or nuclear entry route.
Not much is known about the exact physiological roles of LEDGF/p75. In general, LEDGF/p75 func-

tions as molecular tether for a variety of proteins to the cell chromatin (Figure 1.8) such as the Mixed-Lineage Leukemia/Menin (MLL/MENIN) complex [Yokoyama and Cleary, 2008], CDC7-activator of S phase kinase complex [Hughes et al., 2010], JPO2 [Maertens et al., 2006; Bartholomeeusen et al., 2007] and Pogo transposable element with zinc finger domain (PogZ, [Bartholomeeusen et al., 2009]). Some reports suggested a possible role for LEDGF/p75 as a transcriptional co-activator in promoting cell survival under stress conditions [Shinohara et al., 2002] and functioning in homologous recombination upon induction of a DNA double-stranded break [Daugaard et al., 2012].

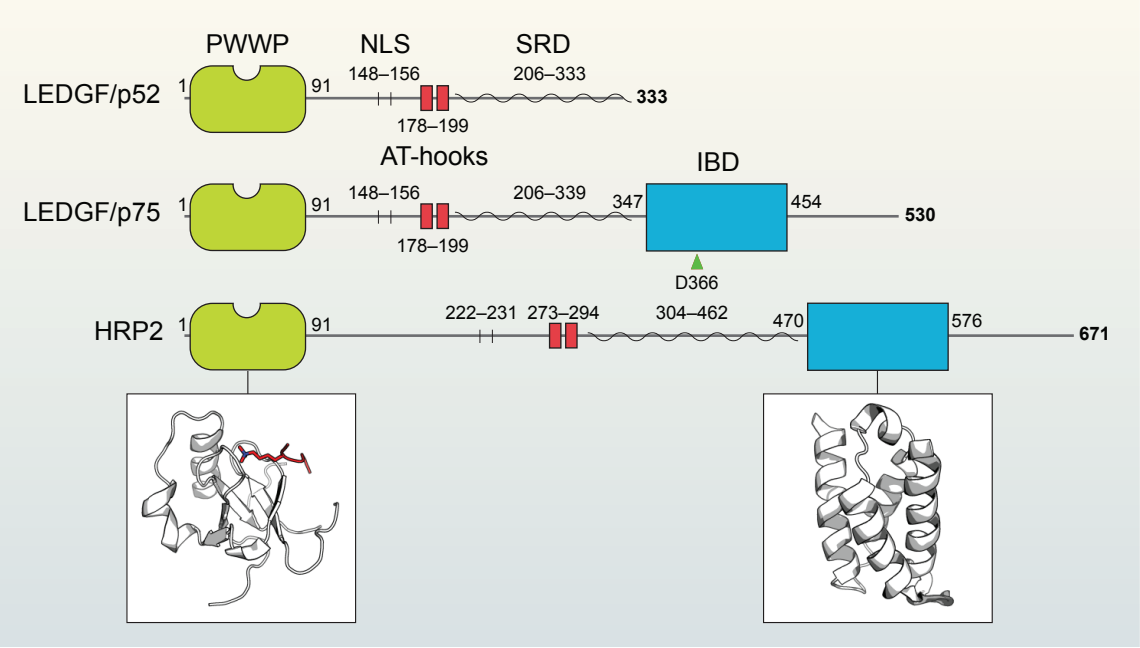


Figure 1.8: Schematic representation of the LEDGF/p52, LEDGF/p75 and HRP-2 domain structure. The *PSIP1* gene encodes for two distinct splice variants LEDGF/p52 and LEDGF/p75. LEDGF/p75 contains a unique C-terminal protein-binding domain, coined Integrase Binding Domain (IBD, blue) responsible for HIV-IN interaction. Several endogenous proteins like JPO2, PogZ and MLL bind to the same interface. At its N-terminal end LEDGF carries multiple chromatin interacting domains; the PWWP domain (green), the AT hook-like domain (AT, red), four charged regions (CR1-4) and a negatively super-coiled DNA recognition motif. D366 is a pivotal amino acid involved in HIV-IN interaction (green arrowhead). Mutation to Asn (D366N) abolishes HIV-IN interaction. HRP-2 is the only human paralogue carrying both a PWWP and an IBD domain. Boxes highlight, from left to right, structures of a PWWP domain bound to a methylated histone peptide (in red), and the IBD. PWWP, Trp- Pro-Pro-Trp; NLS, Nuclear Localization Signal; AT-hook, AT-hook minor groove DNA binding motif; SRD, Super-Coiled DNA recognition region; IBD, Integrase-Binding Domain. *(Modified from [Demeulemeester et al., 2015])*

1.1.6.4 Small molecules targeting the LEDGF/p75-IN interface

Integration into the host cell genome is a pivotal step in the HIV replication cycle and is catalyzed by the viral IN enzyme. Blocking the viral IN therefore forms a valuable therapeutic target for antiviral therapy. Targeting of its catalytic function led to the development of several potent INtegrase Strand Transfer Inhibitors (INSTIs), such as raltegravir (RAL, MK-0518), elvitegravir (EVG, JTK-303/GS-9137) and dolutegravir (DTG, S/1349572) which represent a significant milestone in the field of HIV-1 drug discovery and are currently in clinical use (see Figure 1.10 a on page 17). However, a steady HIV mutation rate together with poor adherence may result in cART escape mutations and the emergence of drug resistant HIV strains causing subsequent treatment failure. Emergence of (cross)-resistance against INSTIs in patients receiving treatment therefore warranted research into novel, allosteric non-catalytic IN inhibitors with distinct mechanisms of action.

LEDGF/p75 knockdown/knockout studies showed that the LEDGF/p75-IN interaction is crucial for HIV replication [Ciuffi et al., 2005; Shun et al., 2007; Marshall et al., 2007; Schrijvers et al., 2012b]. Overexpression of the IBD domain or LEDGF/p75 based peptides outcompeted endogenous LEDGF/p75 and blocked HIV replication [De Rijck et al., 2006; Hayouka et al., 2007]. In 2005 Cherepanov et al. reported the crystal structure of the HIV-IN CCD dimer in complex with the LEDGF/p75 IBD (PDB accession number 2B4J, [Cherepanov et al., 2005]). Amino acids Ile_{365}, Asp_{366}, Phe_{406} and Val_{408} of LEDGF/p75 where shown to be crucial in mediating the interaction with IN. The LEDGF/p75-IN interaction site comprised a well-defined binding pocket and led to the structure based development of two 2-(tert-butoxy)-2-substituted acetic acid derivatives, at the University of Leuven (LEDGIN 1 & 2, depicted in Figure 1.10 b). These compounds comprised a novel class of potent allosteric IN inhibitors and were coined LEDGINs [Christ et al., 2010]. Co-crystals of the identified hit compounds soaked into the HIV-1 IN CCD confirmed the initial pharmacophore binding to the LEDGF/p75 binding pocket on the CCD dimer interface (see Figure 1.9 on page 16), information which was crucial for the design and further development of the more potent congeners CX05045 [Christ et al., 2010] and CX014442 [Christ et al., 2012] with improved biological activities, inhibiting HIV replication in the nanomolar range (Figure 1.10 b). CX014442 inhibits viral replication at an EC_{50} of 69 $\pm$ 3 nM and a selectivity index of 1391 (CC_{50}/EC_{50}) comparable to those of FDA approved drugs. In addition, these congeners displayed a dose-response curve with a steep Hill slope.

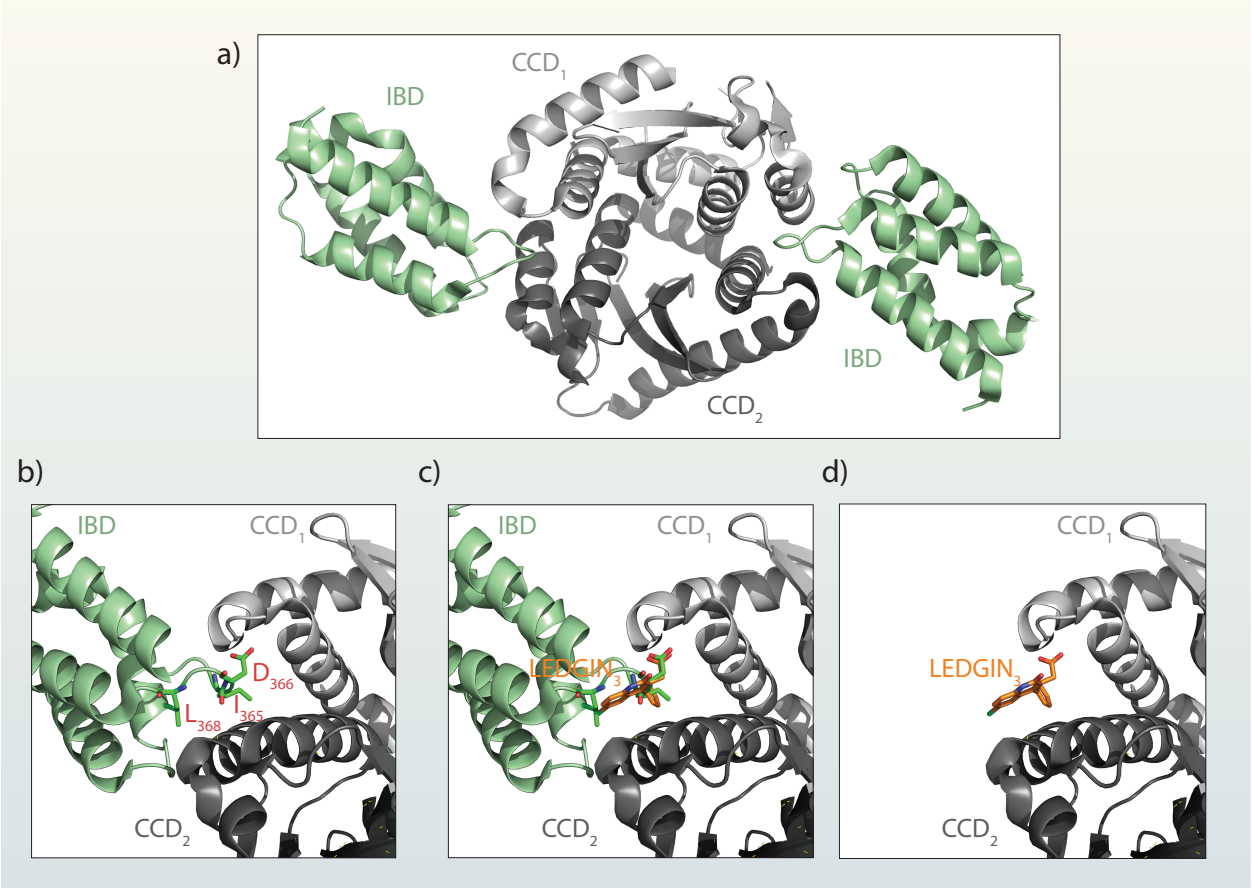


Figure 1.9: LEDGINs bind to the IN dimer interface. a) Structures depicting the LEDGF/p75 IBD (pale green) in complex with a dimer of IN CCDs (light and dark grey) (PDB acession number 2B4J). b) LEDGF/p75 integrase binding domain (IBD) binding to the well-defined pocket formed by two IN CCD dimers. c) Superposition of LEDGIN3 on the LEDGF/p75 IN interface [Christ et al., 2010] (PDB acession number 2B4J and 3LPI respectively). d) LEDGINs are able to displace LEDGF/p75. Their concensus pharmacophores are based on residues Ile_{365}, Asp_{366} and Leu_{368}. Renderings are generated using the PyMOL software [Schrödinger, LLC, 2015].

Apart from structure based design, other groups have independently identified small molecules inhibiting the LEDGF/p75-IN interaction based on a high throughput screening assay, e.g. Boehringer Ingelheim, Canada (BI, Figure 1.10 b [Fader et al., 2014]). The LEDGIN compounds core is relatively solvent exposed and may be varied while the key pharmacophoric feature is an acetic acid group substituted in the 2-position with hydrophobic bulk, preferably ter-but(oxy/yl), paced in ortho from a substituted phenyl ring. The importance of the acetic acid residue lies in mimicking of the LEDGF/p75 residue D_{366}. At present, LEDGINs are still early in the clinical developmental stage (BI 224436) but have gained significant interest of nearly all major pharmaceutical companies involved in the treatment of HIV/AIDS (a patent overview is provided in [Demeulemeester et al., 2014a]).

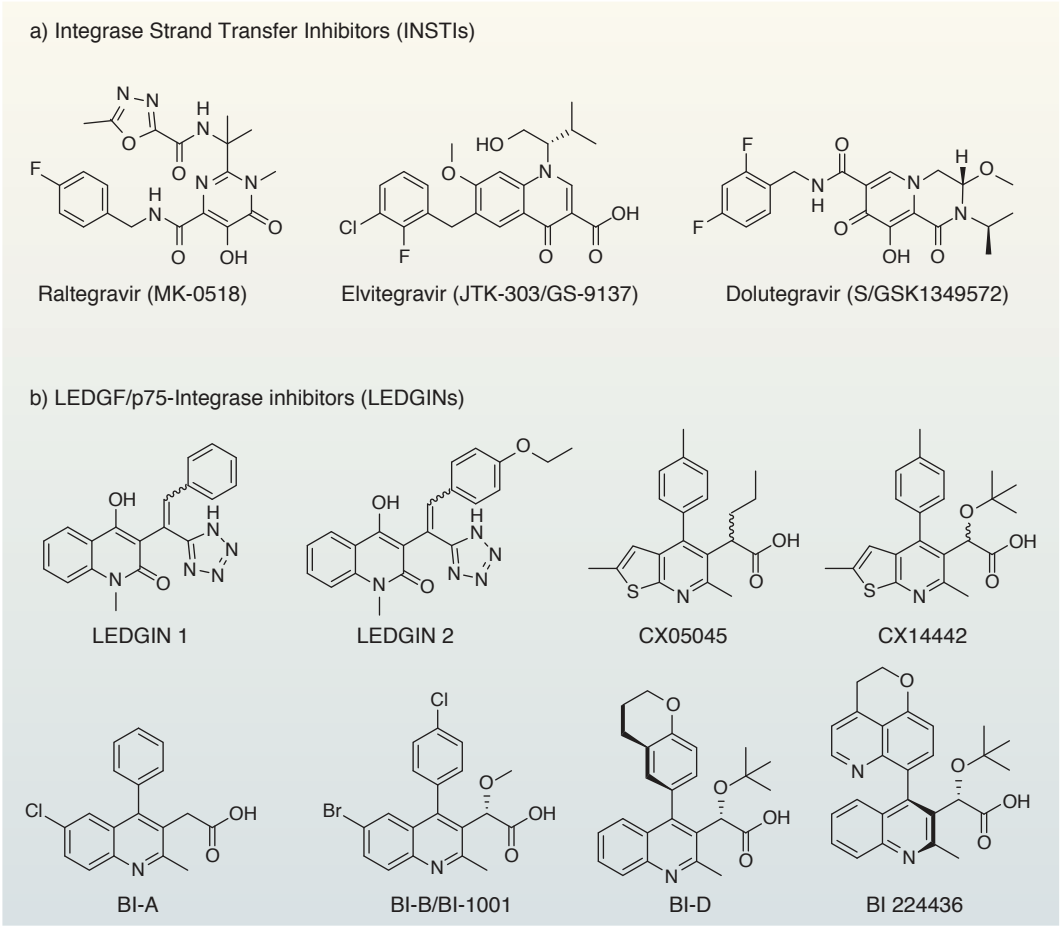


Figure 1.10: INSTIs & LEDGIN congeners. Figure depicting different a) integrase strand transfer inhibitors (INSTIs) and b) LEDGINs developed by the KU Leuven and others (Boehringer Ingelheim (BI) as reviewed in [Demeulemeester et al., 2014a]. *(Modified from [Demeulemeester et al., 2014a])*

LEDGINs display a multi-modal mechanism of action.

By perturbing IN assembly LEDGINs compete with the LEDGF/p75 cofactor and block chromatin tethering of the PIC. Unlike strand transfer inhibitors potent LEDGIN congeners inhibit both strand transfer and 3'-processing catalytic activities, and enhance the formation of IN dimers, disrupting the precise and timely IN oligomeriation which prevents a productive complex assembly ('early effects', Figure 1.11 on page 18, [Christ et al., 2010, 2012; Kessl et al., 2012]). Moreover, binding of LEDGINs to the IN dimer interface is independent of LTR assembly (in contrast to INSTIs). The enhanced potency of LEDGINs in LEDGF/p75 knockout cells highlights the importance of LEDGF/p75 displacement during the early steps [Wang et al., 2012]. Additionally, when LEDGINs are present during the late stages of viral replication the newly produced virions are significantly less infectious showing defects during reverse transcription, nuclear import and integration stages of the next round infection ('late effect', Figure 1.11, [Le Rouzic et al., 2013; Jurado et al., 2013; Desimmie et al., 2013; Gupta et al., 2014]). Both effects synergistically contribute to the overall inhibition of HIV-1 replication and cannot be uncoupled in the infected cell. Transmission electron microscopy studies revealed that roughly 60 % of the produced particles displayed morphological aberrations (ribonucleoprotein mis-localization outside the capsid core or cone malformation). When pre-incubated with HIV IN prior to addition

of the LTR fragment, CX014442 and BI-B/1001 also inhibited 3' processing and strand transfer *in vitro* in the submicromolar range [Christ et al., 2012; Kessl et al., 2012]. Overall, LEDGINs display activity against different HIV clades (HIB_{IIIB}, HIV_{BAL}, HIV_{U2}, $HIV_{NL4.3}$) and clinical isolates [Christ et al., 2010, 2012; Fenwick et al., 2014] but are less active against HIV-2 and SIV due to a amino acid substitution at position IN_{128} involved in the interaction with LEDGF/$p75_{365}$. LEDGIN resistant HIV strains however did not display any cross-resistance to INSTIs. An additive or even modest synergistic inhibitory phenotype is observed when co-administered together with most NRTIs/NNRTIs or INSTIs [Christ et al., 2012; Fenwick et al., 2014].

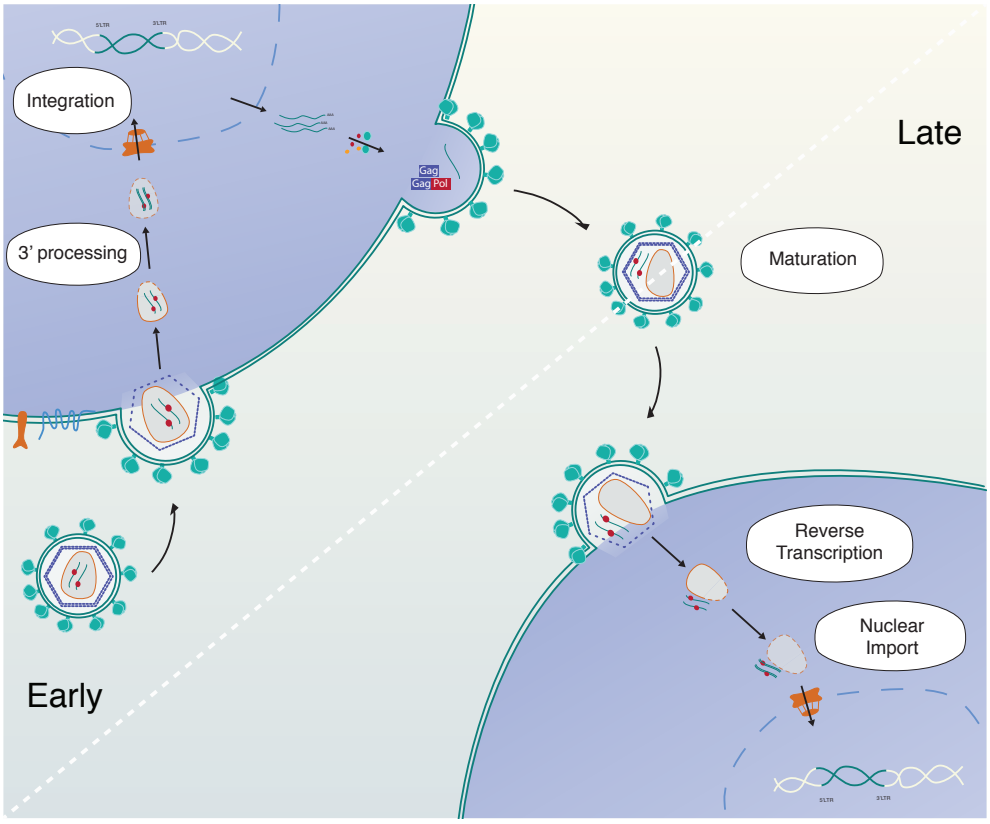


Figure 1.11: Multi-modal mechanism of action of LEDGINs. During EARLY replication steps, LEDGINs act as allosteric inhibitors of HIV IN through a stabilizing effect on IN multimerization, ultimately leading to the inhibition of the 3' processing and DNA strand transfer. Simultaneously the binding of LEDGINs to integrase inhibits the interaction with the chromatin tethering factor LEDGF/p75 leading to a block in HIV integration. When LEDGINs are present during the formation of viral particles, increased formation of IN multimers leads to a maturation defect evidenced by misshaped viral particles with the ribonucleoprotein located outside of the viral core ('late effect'). This maturation defect leads to aborted reverse transcription and defective nuclear import upon *de novo* infection of the host cell.

1.2 Towards an HIV cure

The development of antivirals in the last decades revolutionized the treatment of HIV/AIDS, enabling clinicians to suppress HIV replication to undetectable levels, restore the immune system of an infected individual and profoundly increase a patients survival which turned HIV infection from a life-threatening disease into a chronic disorder. Yet, current antiretroviral therapy fails to cure HIV infection nor does it fully prevent pathology or restore the normal life span of HIV-infected patients. The major impediment towards an HIV cure lies in the existence of a stable reservoir of latently infected cells having a long half-life and carrying replication competent provirus. In addition, anatomical sanctuaries where the drug penetration is poor could allow low levels of ongoing viral replication [Lorenzo-Redondo et al., 2016]. Upon infection of the host cell, HIV can either actively replicate to rapidly produce progeny virions (which typically results in cell death) or enter a long-lived quiescent state. The dormant pool of virus is established in few infected cells early on during acute infection and enables HIV to persist for decades evading host immune surveillance and potent cART. It is the presence of this latent provirus that triggers a rapid rebound of viremia within weeks after therapy cessation requiring patients to adhere to a lifelong suppressive therapy which on its turn may be associated with drug related side effects (reviewed in [Torres and Lewis, 2014]). Moreover, the rapid replication and mutation rate of HIV, its extensive genetic diversity and suboptimal compliance of patients to ART drive HIV in becoming increasingly drug resistant resulting in treatment failure. Therefore, new curative strategies to treat HIV/AIDS are urgently needed and exhaustion of the latent reservoir has become a highly prioritized goal in HIV-1 research. Over the past years, tremendous progress has been made in defining and characterizing this latent reservoir which is summarized in the next subsections.

1.2.1 HIV reservoirs and anatomical sanctuaries

HIV tropism is mainly defined by the interaction of the viral envelope glycoproteins (gp120 and gp41, see section 1.1.5) with the CD4 cellular surface receptor and CXCR4 or CCR5 co-receptors. Hence, HIV specifically targets $CD4^+$ T-cells and cells of the monocyte/macrophage lineage. Other chemokine receptors such as CCR3 and CCR8 can also function as HIV co-receptors and therefore HIV, although to a lesser extent, is also able to infect dendritic cells (DCs), natural killer (NK) cells and some specialized cell types of the central nervous system (CNS) (see Figure 1.12 on page 1.12, [Churchill et al., 2009]). In activated T-cells HIV infection generally proceeds with the generation of progeny virions while in resting T-cells the virus is more likely to enter quiescence as part of the host cell genome. Resting, memory $CD4^+$ T-cells form the most prominent cells contributing to the long-lived latent reservoir, with a frequency of $1/10^6$ resting $CD4^+$ T-cells harboring latent provirus in patients receiving cART (estimated using the viral outgrowth assay, VOA [Eriksson et al., 2013]). Recent experimental evidence however showed that a fraction of the non-reactivated proviruses is still replication competent and that the exact size of the latent reservoir might be larger than initially anticipated. Quantitative PCR measurements suggested a $\pm$ 60-fold underestimation of the reservoir size by VOA measurements [Ho

et al., 2013]). Infection of resting $CD4^+$ T-cells is far less efficient than activated $CD4^+$ T-cells [Swiggard et al., 2005], therefore HIV latency may primarily be established upon transition of activated cells towards a resting memory state [Siliciano et al., 2003]. However, *ex vivo* latency can also be established directly in activated T-cells [Chavez et al., 2015]. Next to the well-established and characterized resting $CD4^+$ T-cell reservoir, several groups have confirmed the presence of latent provirus in the heterogeneity of tissues targeted by HIV including the Gut-Associated Lymphoid Tissues (GALT, [Yukl et al., 2013b; Chun et al., 2008; Yukl et al., 2010]), gentical tract, lymph nodes, CNS and spleen. At present the key question remains which exact cellular source is responsible for reseeding of viral infection upon therapy cessation. More controversial is the question whether ongoing replication or homeostatic proliferation [Chomont et al., 2009; Bosque et al., 2011; Maldarelli et al., 2014] is involved in maintenance of the HIV-1 reservoir. Trace levels of viremia have been previously reported in ART treated patients [Frenkel et al., 2003] supported by genetic divergence [Palmer et al., 2008; Shiu et al., 2009; Fletcher et al., 2014; Lorenzo-Redondo et al., 2016] with impaired drug penetration in lymphatic tissues allowing ongoing replication to contribute to the latent reservoir. Others found an increase in 2-LTR circles following cART intensification using INSTIs [Buzón et al., 2010; Hatano et al., 2013]. These data would point towards a role for ongoing replication in the maintenance of HIV persistence. However, the lack of a reduction of the reservoir size upon cART intensification and a lack of viral evolution in other studies challenged these conclusions [Gandhi et al., 2012; Kearney et al., 2014; Chéret et al., 2015b]. It will become clear in the next paragraphs that both the heterogeneity in cell types and corresponding diversity in molecular mechanisms responsible for the establishment and maintenance of latent provirus together with the lack of biomarkers able to distinguish latently infected cells from uninfected cells complicate the search for an HIV cure.

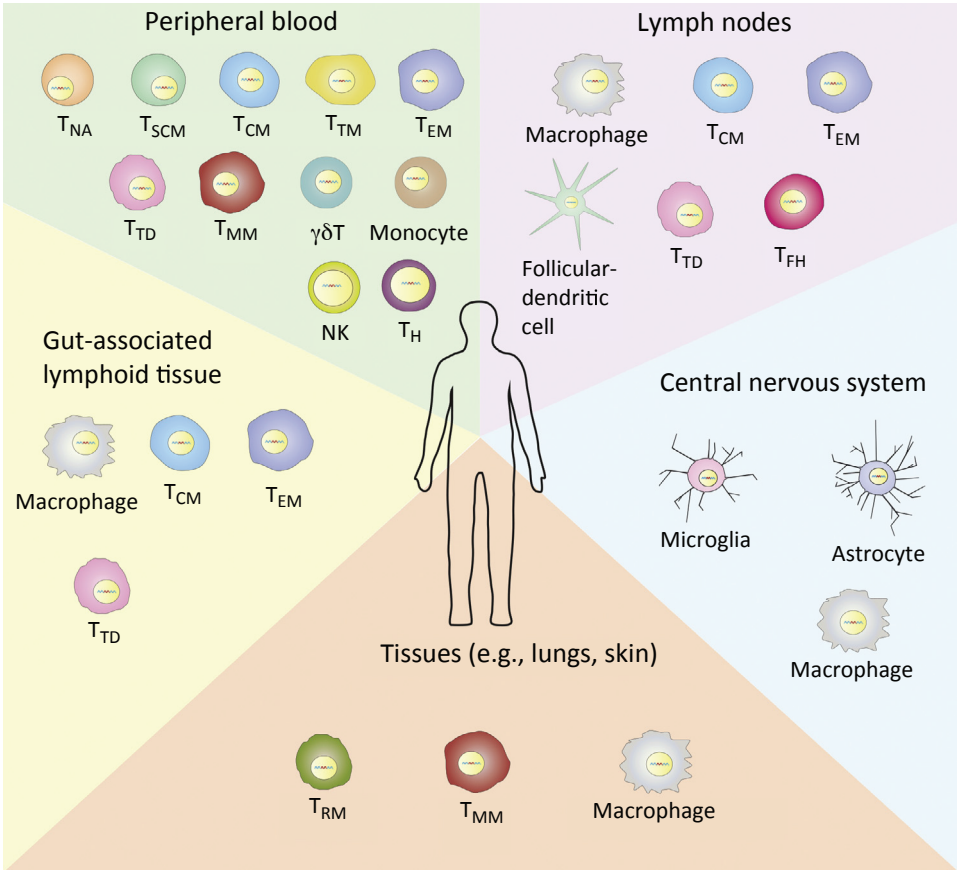


Figure 1.12: Latent reservoirs. Figure depicting the heterogeneity of anatomic focuses harboring latent HIV infections. In addition to resting memory $CD4^+$ T-cell subsets found in the peripheral blood, a significant portion of latently infected cells can be found in the lymphoid tissue, GALT and CNS. Furthermore, tissues such as lungs and skin may also contain latent HIV proviruses and unknown reservoirs remain to be defined. The unique characteristics of each cell type harboring latent provirus illustrates the multifactorial nature of transcriptional latency. *(Figure derived from [Barton et al., 2016])*

1.2.2 Heterogenous regulatory mechanisms of HIV Latency

HIV latency has long been believed to be an epiphenomenon rather than an evolutionary maintained threat until recently some proposed that latency is "hardwired" into the HIV's gene-regulatory circuitry [Razooky et al., 2015; Rouzine et al., 2015]. Modeling studies predicted enhanced transmission of virus having an innate propensity to establish a latent infection which could provide a transmission advantage by enabling the virus to persist in the mucosal environment [Rouzine et al., 2015]. In the following section the major endogenous host-cell mechanisms, host-cell environmental and autonomous viral circuitry controls involved in the regulation of this phenotypic bifurcation are highlighted. At present the exact interplay between these parameters remains largely unclear and may be biologically distinct within different cell populations of the latent reservoir, yielding a number of possible distinct basal transcriptional states for HIV soon after infection. In general, two different types of latency can be designated based on the integration state of the viral DNA; I) pre-integration and II) post-integration latency. Unintegrated HIV-1 DNA (pre-integration latency) will either degrade or integrate in the host cell genome and is likely to present a stable reservoir only in slowly or non-dividing cells such as

macrophages. Post-integration latency on the other hand refers to the presence of integrated retroviral DNA which is transcriptionally inactive. Here, the mechanisms controlling HIV latency can be broadly classified in *cis*-acting such as the site of integration or local chromatin environment and *trans*-acting mechanisms such as transcription factor regulation, cellular relaxation state etc.

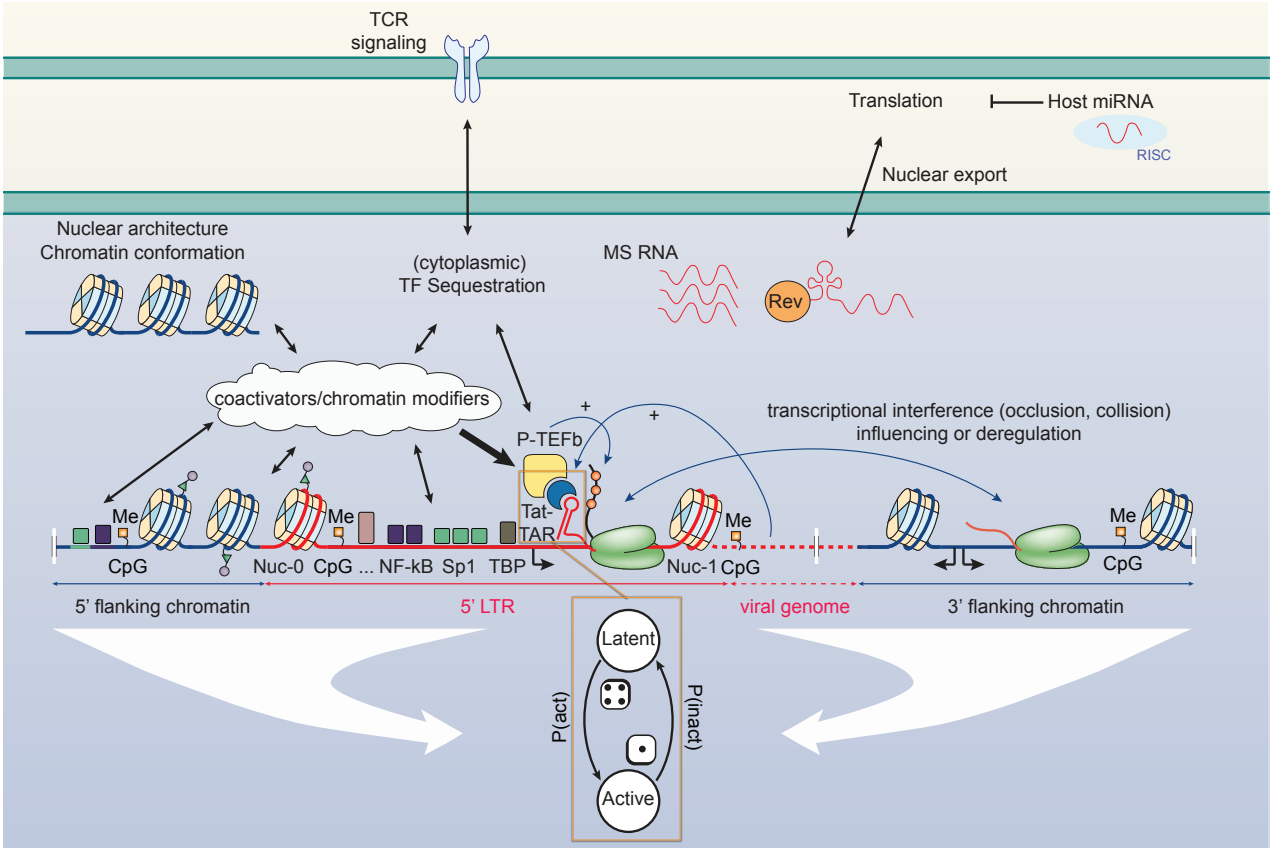


Figure 1.13: Simplified model of HIV-1 proviral transcription. Schematic representation of HIV-1 proviral transcription. The chromatin context, both in *cis* and in *trans*, together with the presence of various LTR-binding transcription factors, determine the intrinsic proviral activity through various other co-activators and chromatin modifiers. The viral Trans-activator of transcription (Tat) binds to the Trans-Activation Response (TAR) RNA stem-loop at the 5' end of viral transcripts and recruits P-TEFb stimulating transcription elongation. Amplification of basal viral expression by this Tat positive feedback creates a robust latency switch. Influencing the different parameters of the switch, results in different probabilities for the ON and OFF states. Me, Methylation; TCR, T-Cell Receptor; MS RNA, Multiply Spliced RNA; miRNA, microRNA; RISC, RNA-Induced Silencing Complex; TF, Transcription Factor; TBP, TATA-Binding Protein; NFκB, Nuclear factor-κB; P-TEFb, Positive Transcription Elongation Factor b. *(Modified from [Demeulemeester et al., 2015])*

1.2.2.1 Viral gene-regulatory circuitry

After integration of the HIV-1 virus in the host cell genome, the virus completely relies on the host RNA polymerase II (Pol II) for RNA transcription and the subsequent generation of viral proteins. This transcriptional process can be generally subdivided in different stages; pre-initiation, initiation, promoter clearance, elongation, and termination [Fuda et al., 2009]. The HIV genome encodes for a transcriptional master circuitry which is driven by the HIV Tat protein. Upon infection, HIV-1

will initially produce only short completely spliced mRNAs encoding the viral regulatory proteins Tat and Rev. Tat or transactivator of transcription protein is an absolute requirement for transcription elongation [Jones, 1997] and will establish a positive feedback loop able to amplify stochastic fluctuations in basal transcription from the viral promoter [Weinberger et al., 2005, 2008; Burnett et al., 2009] by enhancing RNA pol II efficiency and processivity [Peterlin and Price, 2006]. Elongation will result in the synthesis and nuclear export of longer unspliced species of RNA mediated by Rev. These longer RNA products are required for the generation of infectious viral particles. When Tat accumulates above a critical threshold level, HIV transcription elongation is stimulated by binding of the Tat protein to an RNA stem cell loop structure called the *Trans*-Activation Response (TAR) element present at the 5' end of all viral transcripts and recruitment of the Positive Transcription Elongation Factor b (P-TEFb). P-TEFb forms a heterodimer composed of the cyclin-dependent kinase 9 (CDK9) and its regulatory partner cyclin T1 (CycT1) and is able to phosphorylate, either directly or indirectly, the C-terminal domain (CTD) of the largest subunit (RPB1) of Pol II. In addition P-TEFb phosphorylates several negative elongation factors. In the cell, the P-TEFb level is tightly regulated as it is required for elongation of many endogenous genes. P-TEFb can be sequestered in a kinase-inactive complex containing the 7SK small nuclear RNA (snRNA) [Nguyen et al., 2001; Yang et al., 2001]. The integrity of the 7SK snRNP complex is maintained by other associated factors such as HEXamethylene bis-acetamide inducible 1 (HEXIM1), Lupus Antigen Related Protein 7 (LARP7) and MePCE which control sequestration of unused P-TEFb in the inactive state. In resting $CD4^+$ T-cells CycT1 levels are rather low and P-TEFb is sequestered promoting a latent phenotype [Hoque et al., 2011; Chiang and Rice, 2012; Budhiraja et al., 2013]. Under specific conditions including TCR activation or the exposure of cells to hypertrophic or stress signals P-TEFb is released. The free P-TEFb is generally engaged by Brd4, which recognizes specific acetylated histones, and tethered to chromatin regions in order to stimulate transcription elongation. The HIV Tat protein however will directly compete with Brd4 for recruitment of P-TEFb in order to ameliorate HIV transcription elongation (reviewed in [Liu et al., 2014]). Of note, Brd4-P-TEFb appears to be relevant for basal HIV-1 transcription but is inhibitory to Tat-transactivation [Yang et al., 2005]. Recent evidence suggests that stochastic fluctuations in the Tat circuit are sufficient in driving the provirus in one of the respective transcriptional phenotypes [Weinberger et al., 2005, 2008; Burnett et al., 2009]. It appears as if viral evolution evolved to a modular viral circuitry which is able to maintain a remarkable autonomy from environmental stimuli and simultaneously being influenced by probabilistic ON-OFF decision making. The hardwiring of this latency switch has been proposed to represent a bet hedging strategy, preserving fitness in unpredictable, fluctuating environments by probabilistically switch phenotypes optimized to maximize lentiviral transmission.

1.2.2.2 Environmental determinants

Just like expression of endogenous genes, the transcriptional profile of integrated proviral DNA is affected by a combination of environmental parameters. Because eukaryotic chromatin is highly orga-

nized and contains very versatile environments displaying different degrees of compaction it has long been suspected that the integration site and associated local chromatin context could affect the HIV transcriptional phenotype. HIV integration is a non-random process which is preferentially targeted into transcriptionally active regions in proximity to the nuclear rim [Di Primio et al., 2013; Marini et al., 2015]. This integration preference is orchestrated by interaction of the HIV-IN protein with LEDGF/p75 [Ciuffi et al., 2005] and to a lesser extent HRP-2 [Schrijvers et al., 2012b]. Several studies reported on the dependency of the HIV expression levels on the surrounding chromatin context and 3D spatial localization, a variabile inducing up to a 1000 fold difference in expression levels and affecting the transcriptional noise [Jordan et al., 2001; Singh et al., 2010; Dar et al., 2012; Akhtar et al., 2013]. Integration of HIV-1 in heterochromatin regions could repress viral transcription [Jordan et al., 2003; Lewinski et al., 2005] while integration in euchromatin could increase transcriptional activity.

Post-translational modifications of nucleosomal histone tails (such as methylation and acetylation) by chromatin-modifying enzymes play a key role in compaction and amassment of the chromatin and thereby affect the recruitment of available positive and negative transcription factors to the viral LTR promoter. Open chromatin is typically correlated with histone acetylation (regulated by Histone Acetyle Transferases (HAT) and Histone DeACetylases (HDAC)) while condensed chromatin is promoted by histone methylation (by Histone Methyl Transferases (HMT)). Surprisingly, a study involving five-well defined models of HIV latency did not find a significant correlation of specific genomic and/or epigenetic features with a latent phenotype across these models [Sherrill-Mix et al., 2013]. Another variable affecting transcriptional activity is the orientation of the integrated provirus with respect to the host gene. In general, a parallel orientation downstream of an actively transcribed gene increases transcriptional interference by read-through transcription from the RNA polymerase in the host gene [Han et al., 2008; Lenasi et al., 2008] via promoter occlusion while a anti-parallel orientation could lead to convergent transcription and collision of the transcription machinery. However, debate exists as to which extent this proviral orientation affects HIV transcription.

In the transcriptionally inactive state, two key nucleosomes are located at the HIV-LTR. Nuc-(0) spanning the region from position -415 to position -255 and Nuc- (+1) from position +1 to +155 with respect to the transcription start site of HIV genome. Both histones are susceptible to post-translational modifications [Keedy et al., 2009; Friedman et al., 2011]. For several histone deacetylases and histone methyl transferases (EZH2, G9a and SUV39H1) it has been shown that they contribute to latency. The extent by which DNA methylation of cytosine residues within the viral promoter epigenetically regulates HIV transcriptional state remains poorly understood. Methylation of two CpG sites bordering the transcription start site have been linked with transcriptional silencing [Kauder et al., 2009; Blazkova et al., 2009]. Contradictory results, however, have been obtained about the level of LTR promoter methylation *in vivo* [Blazkova et al., 2012; Palacios et al., 2012]. In between the two nucleosomes the LTR promoter contains several recognition elements for different host transcription factors such as NF-$\varkappa$B, NFAT, SP-1, AP-1, LEF-1, COUP-TF, USF, Ets1, and CREB [Verdin et al., 1993; Van Lint et al., 1996; Rohr et al., 2003]. Many of these transcription factors are sequestered to the cytoplasm in resting $CD4^+$ T-cells and will only allow HIV transcription upon T-cell activation. E.g. the NF-$\varkappa$B

heterodimer p65/p50 (active form) is sequestered to the cytoplasm by IκB while the homodimer p50/50 occupies the HIV-1 promoter.

Debate still remains regarding the exact mode of establishment of latent provirus in resting $CD4^+$ T-cells *in vivo*. Several studies showed a cell-driven silencing of HIV transcription related to a change in the cellular relaxation state of activated $CD4^+$ T-cells towards a resting memory state [Finzi et al., 1999; Siliciano and Siliciano, 2004]. While other studies showed that transitioning of primary T-lymphocytes from activated to resting did not silence HIV expression and suggested that the intrinsic viral program was the major determinant controlling the establishment of latency. Based on our current understanding it seems rather likely that the exact combinatorial mechanism will differ between patients, cells and different HIV-strains. The threshold levels of Tat necessary for achieving a productive infection might change over time as cells encounter different stimuli or evolve/ differentiate and age over time.

1.2.3 Current HIV treatment regimens

At present combinatorial AntiRetroviral Therapy (cART) functions as the standard treatment regimen for patients infected with HIV and enables complete suppression of HIV plasma viral loads, dramatically reducing mortality and morbidity. cART treatment radically changed the face of HIV infection from a lethal disease into a manageable chronic condition. A definite cure of HIV infection however, is impeded by the formation, early after primary infection, of viral reservoirs of latently integrated provirus in long-lived blood cells. As a result, HIV is able to rebound viremia within weeks after therapy cessation and patients are committed to a life-long treatment. The financial burden for society together with the treatment burden for patients are significant. In particular, the treatment burden directly affects treatment adherence and imposes a risk on the development of HIV drug resistance.

Currently more than 30 small molecules are approved by the U.S. Food and Drug Administration (FDA) for treatment of HIV infection. These drugs interfere with different steps of the HIV replication cycle (depicted in Figure 1.3 on page 7) and can be classified in 6 different classes: entry and Fusion Inhibitors (FIs, e.g. Maraviroc or Enfuvirtide) inhibit early steps, RT can be inhibited by Nucleoside/Nucleotide Reverse Transcriptase Inhibitors (NRTIs, e.g. AZT) or Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs, e.g. Efavirenz), Protease Inhibitors (PIs, e.g. Darunavir) block proteolytic maturation (late steps). Integrase Strand Transfer Inhibitors (INSTIs, e.g. Raltegravir) block the integration step. cART comprises at least three compounds from two or more classes aiming for synergistic effects and trying to minimize drug resistance development due to a lack in proofreading activity of the reverse transcriptase. Often two nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs, e.g. tenofovir disoproxil fumarate (TDF) or abacavir (ABC)) are combined with a third drug from a different class. A combination with integrase strand transfer inhibitors (INSTIs) is becoming increasingly popular due to their high antiviral potency and excellent tolerability profiles (e.g. Dolutegravir (DTG)) while PI-based treatments display a high genetic barrier to resistance (e.g. Federally approved HIV/AIDS medical practice guidelines; URL: http://aidsinfo.nih.gov/guidelines). In all, the patient's exact regimen is selected based on antiviral activity, potential adverse effects, pill burden, dos-

ing frequency, drug-drug compatibility, resistance barrier, social status, and cost and requires a highly individualized approach with a close monitoring of immune recovery, drug-related side effects and resistance occurrence. Luckily, varying drug pharmacochemical profiles allow alternatives after failure of first-line treatment. New drug combinations aim to relieve the pill burden by providing fixed dose combinations of different drugs as a 'one pill a day' formula (e.g. Atripla, complera, Triumeq, Truvada ...). Increased safety profiles and long-term tolerability together with simplified treatment regimens are of high priority as these improve therapy compliance and decrease the risk of HIV drug resistance. A recent large scale clinical trial on the Strategic timing of AntiRetroviral Treatment (START) [The INSIGHT START Study Group, 2015] concluded to immediately start cART upon diagnosis, independent of the $CD4^+$ T-cell counts. Earlier treatment reduced the transmission rate, reduced the size of the latent reservoir and improved immune recovery and health [Hocqueloux et al., 2013; Le et al., 2013; Tabernilla and Poveda, 2015].

1.2.4 Shock-and-kill approach

Recently, several new therapeutic approaches have been proposed aiming to remit (functional cure) or eliminate (sterilizing/ eradicating cure) the viral reservoir, including novel pharmacological agents, HIV gene therapy, Bone Marrow Transplantation (BMT), immune-modulation and combination approaches. The approach that has become the most prioritized goal within HIV-1 research is the so-called "shock and kill" or "purge and kill" strategy (see Figure 1.14 on page 27). This approach aims to develop therapies capable of exhausting the latent viral reservoir, primarily residing within long-lived resting $CD4^+$ T-cells, by the use of pharmacological agents reverting the HIV transcriptional state by inducing transcription (so-called "shock"). Subsequent production of viral proteins in the latently infected cells should theoretically expose these cells to immune-mediated clearance and/or viral cytopathic effects (so-called "kill"). Uninfected cells at the same time are protected by ongoing cART administration. In general, pharmacologic strategies for reactivation of HIV-1 expression have sought to minimize general T-cell activation avoiding a possible systemic shock.

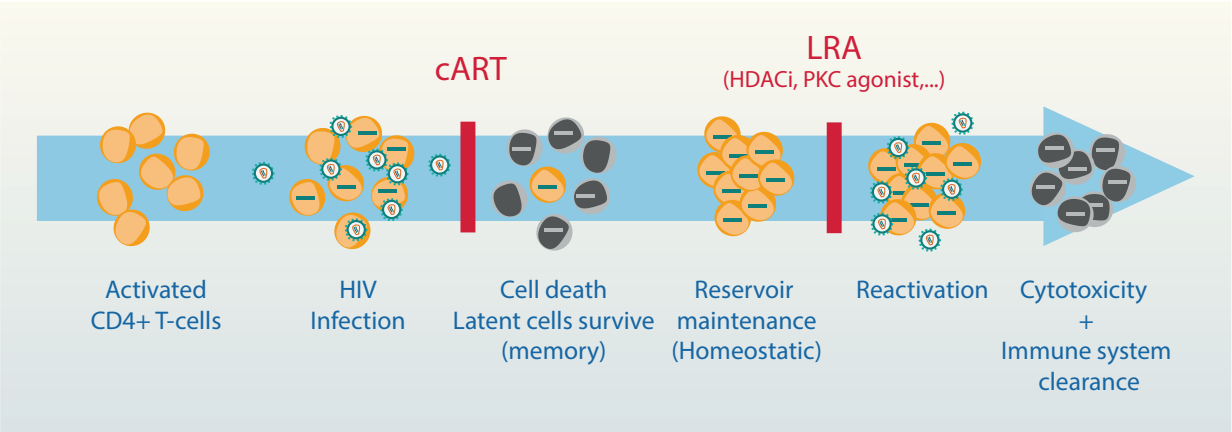


Figure 1.14: Shock and kill approach. The major impediment towards a cure for HIV/AIDS is the existence of a reservoir of latently infected cells able to rebound viremia upon therapy cessation. Stable integration into the host cell genome forever links the fate of the provirus to that of its host. A new therapeutic approach ("shock and kill") aims to reactivate viral expression of dormant HIV in patients. Current treatment regimens such as cART would protect uninfected cells from becoming infected and viral cytophatic effects together with immune system clearing would serve to destroy the infected cells. cART, combination AntiRetroviral Therapy; LRA, Latency Reversing Agent; HDACi, Histone DeACetylase inhibitors; PKC, Protein Kinace C

Histone DeACetylase inhibitors or HDACi have been an attractive choice for pilot eradication trials because they offer an acceptable balance between proviral transcriptional activation and cellular activation. HDACi effects on chromatin remodeling and HIV promoter accessibility have been extensively reviewed [Barton et al., 2014]. The best-studied putative HDACis to date are Vorinostat (SuberoylAnilide Hydroxamic Acid, SAHA), Panobinostat and Romidepsin. Recently, it was shown that HDACis could also promote HIV reactivation by enhancing P-TEFb release [Bartholomeeusen et al., 2013]. Several clinical trials have shown an upregulation of cellular HIV-1 RNA levels often without increases in plasma viremia [Archin et al., 2012, 2014a; Rasmussen et al., 2014; Søgaard et al., 2015]. Hence, the modest outcomes of these clinical trials have led to a re-evaluation of these agents as lead compounds. It remains unclear whether increases in cellular HIV transcription are the result of cellular read-through transcripts and correlate with functional unspliced HIV RNA or translation of viral proteins and subsequent detection by the immune system.

Next to HDACi, a wide range of Latency-Reversing Agents (LRAs) has been investigated *in vitro* and *ex vivo* [Shang et al., 2015; Darcis et al., 2015] acting on different pharmacological targets but few compounds however advanced into clinical trials and none have shown any durable effect on the size of the latent HIV-1 reservoir [Elliott et al., 2014; Rasmussen et al., 2014; Søgaard et al., 2015]. This in part could be attributed to the variety of cell types being part of the persistent HIV reservoir and the heterogeneity of molecular mechanisms involved in regulation of the HIV transcriptional state. A tremendous effort went into the development of *in vitro* models of HIV latency (compared and reviewed in [Archin et al., 2014b; Spina et al., 2013]) but none of these precisely reflect the multifactorial nature of cells harbouring latent provirus from patients.

Protein Kinase C (PKC) agonists (such as phorbol esters, prostratin, bryostatin-1 and ingenol derivatives) appear to be some of the most potent reactivating agents across different models, exerting

their action by promoting the NF-κB pathway. Degrading IκB triggers a translocation of the active NF-κB heterodimer to the nucleus promoting transcription [Trushin et al., 2005]. In addition PKC agonists stimulate production and release of P-TEFb [Sung and Rice, 2006; Fujinaga et al., 2012; Pandeló José et al., 2014]. This bipolar profile makes PKC agonists interesting molecules to purge the latent reservoirs. Moreover, PKC agonists also downregulate the expression of the HIV-1 receptor CD4 and co-receptors CXCR4 and CCR5. Other examples of LRAs focus on HMTi, interference with the BRD4-P-TEFb competition (JQ1) or DNA methylation inhibitors (DNAMTi, 5-AzadC).

Although it is clear that individual LRAs can trigger the production and release of virions from a subset of infected cells *in vitro*, more potent LRAs are required. Therefore, current efforts opt for using combinatorial approaches as the therapeutic potential was illustrated by strong synergy profiles observed when co-administration of two kinds of LRA classes working on distinct regulatory mechanisms [Bouchat et al., 2012; Laird et al., 2015; Darcis et al., 2015].

In the face of the challenges presented by viral latency, measurable and incremental progress has been made. Early eradication studies provided vital information on the nature of the latent reservoir. Important questions remain unanswered with regard to the ideal combination of agents necessary to perturb the heterogeneous reservoir and result in cell death or enable immune recognition. Debate remains as to what are the most reliable means of quantifying latency reversal and reservoir depletion. To completely disrupt HIV latency it will be of key importance to also better understand HIV mRNA export, splicing, translation, antigen expression/processing and presentation allowing latently infected cells to be revealed to or targeted by the immune system. Of note, latent viruses often contain cytotoxic T-lymphocyte (CTL) escape mutations which might, despite successful reversion of latency, impair immune clearance and require CTL stimulation [Shan et al., 2012; Deng et al., 2015].

1.2.5 Alternative approaches to cure HIV

The strategies to achieve an HIV-cure can be broadly classified into three categories: a complete eradication (eliminating all viral reservoirs), a functional cure (immune control without reservoir eradication) and a combinatorial or hybrid cure (reduced or altered reservoir with augmented immune control). These approaches taken towards an HIV cure involve (i) host cell modification to restrict HIV-1 infection, (ii) T-cell engineering to specifically recognize HIV infected cells, (iii) broadly neutralizing antibodies, (iv) therapeutic vaccination, (v) latency enforcement and (vi) the afore mentioned shock-and-kill approach.

To date the only documented case of an HIV eradication is Timothy Ray Brown, referred to as the so-called 'Berlin patient' [Hütter et al., 2009; Yukl et al., 2013a] who was HIV-1 positive and diagnosed with acute myeloid leukemia. Timothy Ray Brown experienced a near complete immune system replacement after myeloablative conditioning, through allogenic Hematopoietic Stem Cell Transplantation (HSCT) from a donor carrying a homozygous *CCR5 δ32* mutation. The mutated CCR5 co-receptor leaves $CD4^+$ T-cells resistant to most HIV-1 strains. This approach, unfortunately, is not broadly applicable due to its high morbidity and mortality rate. In addition, only a low percentage of compatible donors

carry this mutation. Therefore, the approach is only suitable for those with life-threatening hematologic malignancies. Of Note, at present at least another six patients received a graft from a donor screened for CCR5 δ32/δ32 homozygosity. However, none of these patients survived for longer than 1 year. As reported by Kordelas et al. this was explained due to the rebound of a CXCR4-tropic HIV-1 variant [Kordelas et al., 2014]. Among four of the six patients the grafts were associated with a higher rate of early death. Continued administration of cART until stable donor chimerism and appropriate HLA matching CCR5 δ32/δ32 are worth careful consideration [Verheyen et al., 2014]. Hence, most approaches are focusing on achieving a functional or combinatorial cure.

With the advent of gene editing technologies such as Zinc-finger nucleases, Transcription activator like effector nucleases (TALENs) and Clustered Regularly-Interspaced Short Palindromic Repeats-Cas (CRISPR-Cas), autologous $CD4^+$ T-cells or HSCs have been engineered to constrain HIV infection. By generating a knock out for e.g. the CCR5 gene the ability of HIV-1 to enter via its co-receptor can be blocked [Holt et al., 2010; Tebas et al., 2014; Li et al., 2013]. This approach recently advanced to clinical trials. It is however still unclear to what proportion cells need to be modified to constrain HIV replication and whether CXCR4-tropic virus emerges over time. Others are trying to directly target the provirus disrupting it using a CRISPR-Cas based platform [Liao et al., 2015]. Off-note, measures are to be taken to decrease the intrinsic propensity of gene editing technologies for causing off-target effects (reviewed in [Stella and Montoya, 2015]). Alternatively, designer immune responses are gaining interest, having proven their efficacy in combating B-cell lymphomas [Porter et al., 2011]. An example is the generation of Chimeric Antigen Receptor (CAR) expressing $CD8^+$ T-lymphocytes targeting different invariant HIV epitopes. The major concern however is related to possible off-target cross-reactivity [Linette et al., 2013]. T-cell mediated immunity is an event observed in a rarity group of HIV positive patients called 'elite controllers' capable of suppressing viral replication in the absence of cART. Understanding of the mechanisms behind their effective $CD8^+$ T-cell responses and occurrence of favorable HLA alleles would contribute to the development of designer immune responses and proper therapeutic vaccination. Another encouraging report on immune control is the occurrence of post-treatment control in a subset of individuals in the VISCONTI cohort [Sáez-Cirión et al., 2013]. The immunologic correlates however are yet to be defined. A valid alternative approach is based on the recently discovered broadly neutralizing antibodies (bNAbs). bNAbs hold great promise for future HIV-1 treatment but it will remain challenging to elicit proper HIV-specific bNAb-like antibody responses through therapeutic vaccination due to the required extensive affinity maturation [Mascola and Haynes, 2013]. On the contrary, passive infusion of a cocktail of non-autologous bNAbs could form a valid alternative. Here the risk of triggering anti-bNAb antibodies complicates their applicability. It will be vital for vaccine researchers to better understand and overcome the challenges with glycan shielding of conserved epitopes (a review on HIV-1 neutralizing Abs can be found in [Mascola and Haynes, 2013]). The latest of the approaches taken towards an HIV cure has been the most actively pursued (shock-and-kill), however at present short term treatment using a single LRA dose is currently unable to reduce the size of the latent HIV-1 reservoir. Clinical trials have shown significant increases in viral transcription levels but it remains unclear whether the *in vivo* viral reservoirs can be depleted either by

viral cythopatic effects or by immune system clearance. These clinical outcomes have led to the initiation of new trials based on a combination of different LRAs together with therapeutic HIV-1 vaccination or immune therapy boosting of the existing or novel immune responses to recognize and clear infected cells. In contrast to the "shock and kill" approach several groups have advocated for the enforcement of a permanently suppressed transcriptional state. With endogenous retroviruses constituting roughly 8 percent of the human genome being permanently suppressed and nonpathogenic [Jern and Coffin, 2008], mimicking of such a permanently silenced endogenized state may provide an attractive strategy for achieving an HIV remission manageable by the immune system. A rather controversial idea which until now has not been extensively pursued. Exciting progress has been made regarding the suppression of the Tat-based positive feedback loop which drives HIV transcription recruiting P-TEFb [Jin et al., 2016].

1.2.6 Conclusion

Whereas significant progress has been made in the treatment of HIV/AIDS during the past decades and cART is saving the lives of millions of people across the world, a complete HIV cure is still lacking. It is clear that much remains to be learned about the multi-factorial and probabilistic nature of the latent reservoir. Although the pursuit of the "shock and kill" approach provides us with significant novel insights in the biological mechanisms by which HIV operates, clinical outcomes have been rather modest. Moving beyond a complete focus on this strategy, exploring alternative approaches seems rather mandatory. Careful consideration must be given to the ethics of translational research with volunteers infected with HIV living a rather healthy life. In addition, reducing the pill and financial burden, possible drug-related side effects and improving world wide accessibility of cART treatment should be of absolute priority, necessary to increase compliance and reduce the risk of HIV drug resistance.

1.3 Lentiviral vectors for safer gene therapy

In general gene therapy encompasses all those therapies that interfere with a patients genetic material at the genomic, transcriptional or the translational level, a concept that has long appealed to biomedical researchers and clinicians as it promised to treat nearly all diseases at their origins. This therapeutic approach can be roughly classified in three categories; gene addition, gene knockdown/ knockout and gene correction/ alteration or editing of which gene addition is the most commonly pursued. Here, treatment is achieved by counteracting or replacing a malfunctioning gene within cells being adversely affected by the condition (Figure 1.15 on page 31), an approach which is mostly appealing to monogenic, autosomal recessive disorders. During the past decade gene therapy based on stable insertion of retroviral or lentiviral vectors has evolved towards a reproducible cutting-edge technology and an effective treatment option for several hematological malignancies, mainly Primary Immunodeficiency Disorders (PIDs, [Rothe et al., 2014a; Touzot et al., 2015]). Retroviral vectors are excellent tools to stably insert a therapeutic gene into the host cell genome, thereby modifying its genetic characteristics indefinitely and avoiding diminution of the therapeutic effect by dilution. The efficiency by which cells can be genetically corrected and the large therapeutic benefit has shifted gene therapy towards a clinical reality. In a recent retrospective study a gene therapy treatment even outperformed an allogenic haploidentical Hematopoietic Stem Cell Transplantation (HSCT) accounting for a faster immune reconstitution and improved thymus response [Touzot et al., 2015]. Recently a first retroviral vector (Strimvelis^{TM}) was approved by the European Commission for the treatment of ADA-SCID [Cicalese et al., 2016], a leap forward in the gene therapy field.

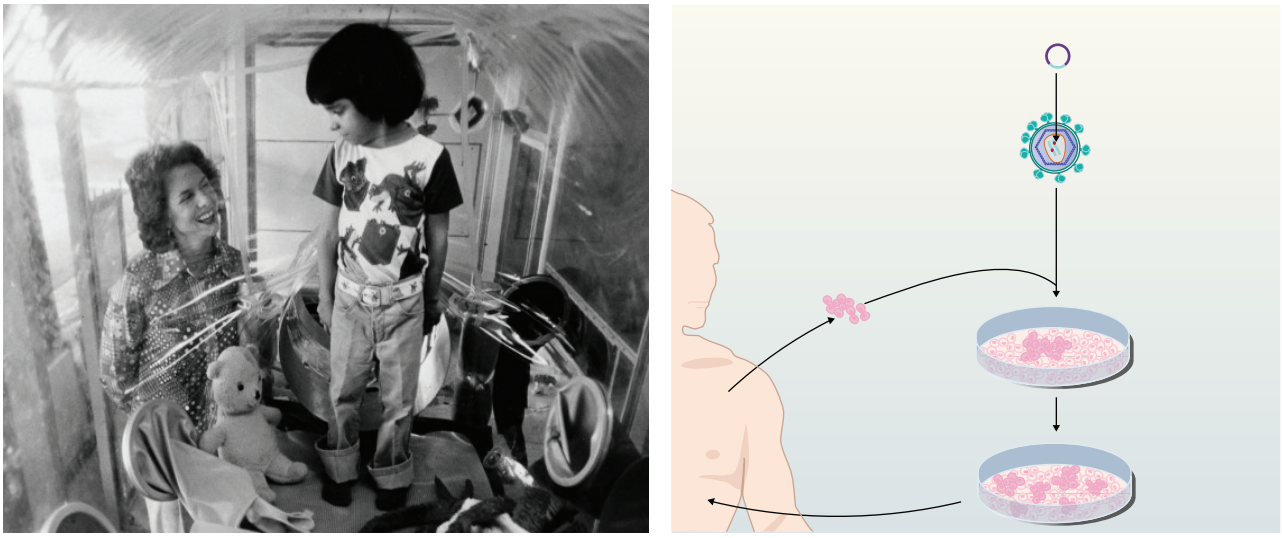


Figure 1.15: Gene Therapy. David Vetter "The Bubble Boy" (1971 - 1984) had a rare monogenic immune disease, X-linked Severe Combined Immune Deficiency (X-SCID), that required him to live in a sterile environment to protect him from infections. If born today, Vetter would likely have been a perfect candidate for an alternative gene therapy treatment able to provide him with a functional immune system.

1.3.1 Re-harnessing viruses to generate viral vectors for efficient gene transfer

Over the past million years, viruses have evolved ingenious ways to efficiently transfer their genetic material into a variety of cell types as they infect a host cell and subsequently hijack the host cell machinery for progeny virion production that in turn are able to infect other cells (see section 1.1.5 on page 6). Viral vectors, on the other hand, transduce a host cell and once stably integrated in the host cell chromatin, will not replicate but employ the cellular machinery to produce the therapeutic gene product. Over the past 30 years, viral vectors have been engineered to be efficient gene transfer tools which are intrinsically replication deficient based on the segregation of genetic information over different plasmids for *cis*- and *trans*-acting elements (see Figure 1.16 b - d on page 32). Genes for *trans*-elements encode structural proteins. This segregation over different plasmids enables the transfer plasmid (Figure 1.16 c) to hold a gene of interest (GOI) and reduces the possibility of forming replication-competent virus via homologous recombination events. Transfection of all constructs in producer cells will result in the generation of vector particles that will transfer the gene of interest into the target cell.

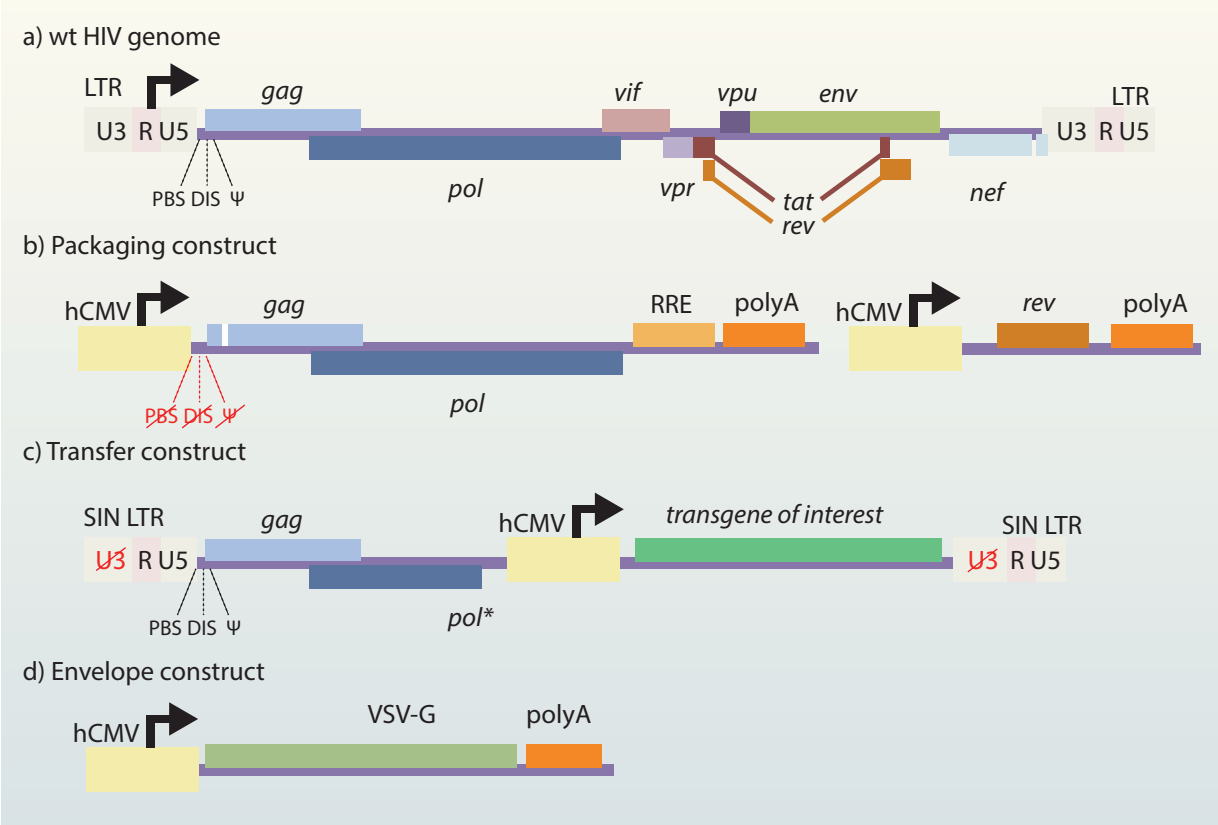


Figure 1.16: Plasmid constructs required for efficient viral vector production. a) Genetic organization of WT HIV virus b) Packaging constructs provide mainly structural viral proteins in *trans* required for vector production c) Transfer construct contains the packaging signal and encodes the gene of interest expression cassette driven by a constitutively active promoter, flanked by inactivated LTR ends. d) Envelope construct allows for pseudotyping of viral vectors by providing an heterologous envelope gene. ψ, packaging signal; PBS, Primer Binding Site; DIS, genome Dimerization Initiation Signal; LTR, Long Terminal Repeat; hCMV, human CytoMegalo-Virus promoter; VSV-G, Vesicular Stomatitis Virus Glycoprotein G.

1.3.2 Severe adverse events and the risk of insertional mutagenesis

Already in the early 1970's, gene therapy was predicted to be the next-generation treatment option for a plethora of genetic disorders. In the early nineties, Adenosine DeAminase Severe Combined ImmunoDeficiency (ADA-SCID) was treated successfully using MLV-based vectors [Bordignon et al., 1995; Blaese et al., 1995; Onodera et al., 1998], fueling the development of additional gene therapeutic approaches for other monogenetic immunedisorders such as X-SCID. In the late nineties, several other patients were treated successfully, providing a functional immune system to these so-called 'bubble'-boys (see Figure 1.15 on page 31) that were sentenced to a life in a sterile bubble. In 2003, however, a major setback tempered the broad application and general acceptance of gene therapy as a valuable treatment alternative [Hacein-Bey-Abina et al., 2003a]. Initial reports of Severe Adverse Events (SAE) in X-linked Severe Combined Immunodeficiency (SCID-X1,[Hacein-Bey-Abina et al., 2003a, 2008]) and Chronic Granulomatous Disease (CGD, [Ott et al., 2006; Stein et al., 2010]) gene therapy trials rose concern and scepticism over the further deployment of such strategies as an alternative to haematopoetic stem cell (HSC) transplantations, but provided vital information to improve the safety of future gene therapy trials. It became apparent that retroviral vectors exhibited common integration sites and that the observation of aberrant T-cell proliferation was associated with the integration of a viral vector in proximity of proto-oncogenes (e.g. SPAG6, CCND2, LMO2 [Ott et al., 2006; Stein et al., 2010]). The presence of strong enhancer sequences in the viral LTR promoter of the early generation Moloney murine Leukemia Virus (MLV)-derived gammaretroviral viral vectors triggered transcriptional dysregulation of neighbouring genes [Hacein-Bey-Abina et al., 2003b; Ott et al., 2006; Howe et al., 2008; Stein et al., 2010]. Later, massive parallel integration site sequencing allowed detailed integrome studies, that highlighted the intrinsic integration preferences of different retroviral vector platforms but were not fully capable of explaining the vector-related SAEs [Aiuti et al., 2007]. The generation of a plethora of clinical and preclinical safety data resulted in important advancements in the biosafety of gene transfer technologies. Development of Self-INactivating Long Terminal Repeats (SIN LTRs, see Figure 1.16 c) lacking strong enhancer sequences and a shifted interest to lentivirus-derived viral vectors substantially reduced the insertional genotoxicity observed in X-linked AdrenoLeukoDystrophy (ALD, [Cartier et al., 2009]) and β-thalassemia patients ([Cavazzana-Calvo et al., 2010]).

However, at present no definitive consensus exists on the real combinatorial causes of the adverse events observed in the distinct clinical contexts. A retrospective study pin-pointed the complex interplay of different parameters such as vector platform used, vector design and promoter, vector insertion profile [Cartier et al., 2009; Cavazzana-Calvo et al., 2010], mean viral copy number, viral splice sites [Trono, 2012], disease background [Shou et al., 2006; Ginn et al., 2010; Bosticardo et al., 2009], transgene nature and transgene product levels [Woods et al., 2006; Grez et al., 2011] having a possible influence on the likelihood of insertional genotoxicity manifestation in a trial [Rothe et al., 2014b]. So far, current assays to test the mutagenic potential of integrating viral vectors lack substantial power and capability to fully predict the exact levels of toxicity that might be observed in any given clinical gene therapy context and no standardized scoring for the genotoxic potential of a new viral vector exists [Rothe et al., 2014b].

The next subsections focus on the effect of the vector integration profile and the search for genomic safe harbor regions.

1.3.3 Towards a safer retroviral integration profile

Retroviral integration was long time considered to be a random process, with the viral IN acting as the major determinant. Large scale integration site sequencing studies revealed a genus specific integration preferences with lentiviruses and gammaretroviruses integrating in the proximity of genes [Schroder et al., 2002; Mitchell et al., 2004]. Lentiviruses displayed an integration pattern that did not show a preference for strong enhancer and promoter regions. Next to improvements in vector design, initial efforts to further lower the toxicity profile aimed at targeting integration away from active transcription units. With the viral IN being the prime determinant of the integration site itself, initial efforts to retarget lentiviral integration were based on fusions between IN and DNA binding domains (e.g. the *E. coli* LexA repressor [Holmes-Son, 2002], bacteriophage λ-repressor [Bushman, 1994; Bushman and Miller, 1997] or engineered zinc finger proteins (Zif268) [Tan et al., 2006]) of which most retained catalytic activity *in vitro* but often significantly impaired vector production upon incorporation in vector particles and additionally decreased transduction efficiency. When it became clear that retroviruses co-opted different cellular proteins to tether their PICs towards the cellular chromatin several groups aimed at modifying this interaction. In the case of lentiviruses, LEDGF/p75 is the major cellular cofactor orchestrating integration [Ciuffi et al., 2005]. Replacement of its PWWP chromatin binding domain with alternative DNA binding modules resulted in an increased integration near features recognized by those domains [Ferris et al., 2010; Silvers et al., 2010; Gijsbers et al., 2009]. LEDGF-fusions were shown to rescue a X-CGD disease phenotype in cell culture [Vets et al., 2013]. Only recently Bromodomain and ExtraTerminal domain containing family of proteins (BET proteins; BRD2, BRD3, and BRD4 [Gupta et al., 2013; Sharma et al., 2013; De Rijck et al., 2013]) were reported to guide gammaretroviral integration by direct interaction with the viral IN. A single IN_{W390A} point mutation [El Ashkar et al., 2014] abrogated the interaction and resulted in a reduced integration in the neighbourhood of promoters and regulatory elements. Next to retargeting of integration for the existing MLV or LV platforms newer vector systems are being explored because of their closer-to-random integration pattern such as alpharetroviral vectors [Suerth et al., 2012; Moiani et al., 2014] or transposon based platforms [Moldt et al., 2011; Turchiano et al., 2014] which could possibly reduce the probability of SAEs. Off note, approaches where stable genetic modification of the host genome is required will always carry an intrinsic risk of genomic perturbation to be related to the scale of the correction.

1.3.3.1 Safe harbour regions

Since the clinical observation of the retroviral integration site being a responsible for diminished therapeutic effects or dysregulation of endogenous genes leading to malignant host cell transformation [Hacein-Bey-Abina et al., 2003b; Ellis, 2005; Cavazzana-Calvo et al., 2010], several strategies have

aimed for redirecting integration towards distinct regions. In parallel, different criteria were put forward in order to define an "ultimate" integration environment able to accommodate integration of new genetic material in a manner that enables functionality prediction and does not harm the host cell or organism. Such regions are referred to as "Genomic Safe Harbor" regions (GSHs) and should minimally perturbate or dysregulate the transcriptional profile of neighbouring genes. The basic criteria aim to exclude regions in close proximity to transcription start sites (<50 kb), oncogenes (<300 kb) or miRNA coding regions (<300 kb), transcription units and ultraconserved elements (defining potentially unsafe integration events, [Papapetrou et al., 2011; Sadelain et al., 2011a; Papapetrou and Schambach, 2016]). Three different loci have been mostly targeted for transgene insertion in the past: (**i**) a common integration site of the human non-pathogenic AAV, located between exon 1 and intron 1 of the protein phosphatase 1 regulatory subunit 12C (*PPP1R12C*) on chromosome 19, known as the AAV site 1 (AAVS1); (**ii**) the chemokine (C-Cmotif) receptor 5 (*CCR5*) gene, a chemokine receptor gene known as an HIV-1 co-receptor for which a homozygous deletion is found in seemingly healthy individuals; and (**iii**) the human orthologue of the mouse Rosa26 locus, a locus extensively validated in the murine setting for insertion of ubiquitously expressed transgenes [Liu et al., 1996; Kotin et al., 1992; Irion et al., 2007; Perez et al., 2008]. Even though these loci do not perfectly match the above mentioned criteria these sites gained interest due to the functional data available on expression robustness or transcriptional dysregulation [Lombardo et al., 2011] and may therefore be of interest for research applications but require a higher burden of proof of safety for clinical applications. Recent evidence indicated that the AAVS1 locus does not faithfully support expression of all transgenes in all cell types and might be silenced by DNA methylation [Ordovás et al., 2015]. Moreover, information is lacking about the effect of *PPP1R12C* haploinsufficiency in different cell types. CCR5 KO increased susceptibility to disease caused by the West Nile Virus [Glass et al., 2006] and Japanese Encephalitis Virus [Larena et al., 2012]. Whereas previously GSHs were mostly empirically discovered or defined on the basis criteria and known genome annotations, it became clear that the definition of a GSH had to be amended to take into account the rapid discovery of new genomic elements (e.g. non-coding RNA or ncRNA), epigenetic modifications (e.g DNA/histone tail modifications, nucleosomal remodeling), spatial organization (e.g. Topologically Associated Domains (TADs))[Papapetrou and Schambach, 2016] and long range interactions. At present, no chromosomal location in the human genome has been demonstrated to qualify as a *bona fide* GSH and the validation of a GSH will highly depend on reciprocal interactions between a transgene and the cell's genomic context. GSHs should maintain "neutrality" in the context of both the linear genome as well as for longe-range interactions. Whereas bioinformatic predictions can be a useful first step, future GSHs will need to be functionally validated in relevant settings since expression and genotoxicity might be cell or differentiation stage dependent. The discovery of potential GSHs, in parallel, will require increased efficiencies in site-specific gene targeting supporting near WT transduction efficiencies (e.g. CRISPR-cas9, zinc-finger nucleases, meganucleases, TALE nucleases).

1.3.3.2 Pre-clinical assays to assess safety

Before new gene therapy vectors enter clinical phases suitable preclinical vector toxicity data need to be provided and approved by the EMA (Europe), FDA (US) or TGA (Australia). The current *in vitro* and *in vivo* assays however lack the potency to completely predict a possible manifestation of SAEs. In the hematopoietic system, a murine bone marrow transplantation model is used where lineage negative HSCs or tumor prone HSCs (Cdkn2a$^{-/-}$ FVB/N.129 mice) are transduced *ex vivo* and transplanted in wt or disease specific mice [Stein et al., 2013] allowing for a longer follow up period of the tempo-spatial integrome [Montini and Cesana, 2012]. Neoplasia and integration near proto-oncogenes or high risk insertions are monitored and the clonal size quantified. In addition a histopathological analysis is performed and a biospatial profile obtained of the genetically modified cells. In conjunction several *in vitro* assays allow for the analysis of vector induced differences in the cellular transcriptome or ability to transform mouse lineage negative cells *in vitro* (*In Vitro* Immortalization Assay, IVIM assay). Here lineage negative cells are isolated (C57BL6/J mice) transduced at using high vector titers, replated at low cell density in the abscence of cytokines. Colony formation accounts for insertional mutagenesis and vector genotoxicity. Uncertainty about the predictive capability of these assays and the limit of detecting parameters contribution to a safer genotoxicity profile together with a cell specific readout indicate the need for improved, more general and/or cell-type/ disease specific assays to evaluate vector safety.

1.3.4 Conclusion

Gene therapy has emerged as an attractive alternative strategy to successfully treat a series of monogenic recessive disorders and enables to circumvent immune conflicts in the setting of allogenic cell transplantation when HLA-identical siblings are lacking. In a recent study a gene therapy treatment even outperformed the haploidentical HSCT [Touzot et al., 2015] accounting for a faster immune reconstitution. While gene therapy often significantly prolongs the lifespan of patients and increases the quality of life, several clinical challenges are yet to be overcome. Important advancements have been made regarding the safety of gene transfer technologies but no consensus has been reached on the exact and intrinsically different combinatorial causes of the adverse events observed in distinct clinical contexts. Sophisticated design of gene transfer vehicles, a better understanding of the interplay with the local chromatin environment together with more sensitive and predictive tests able to predict the therapeutic outcome will allow for gene therapy to become a widely accepted treatment option.

2

Objectives

After decades of antiretroviral therapy, complete eradication of the HIV virus (HIV cure) remains an elusive goal due to the existence of stable reservoir of latently infected cells. Therefore leading scientists shifted strategies to exhaust the latent reservoir in their aim for a functional cure allowing at least interruption of chronic cART administration. Stable integration is a hallmark of retroviruses and a non-random process displaying genus specific integration site preferences. In the case of lentiviruses LEDGF/p75 functions as the major cellular co-factor orchestrating integration towards transcriptionally active regions. So far the evolutionary advantage of these different integration profiles remains incompletely understood. As the surrounding chromatin environment of the integrated provirus is generally believed to affect the transcriptional activity of the virus I studied the role of proviral integration site positioning in HIV persistence. Likewise, the retroviral integration allows for permanent modification of its host cell and makes retroviral vectors interesting tools for gene therapeutic applications. However, the intrinsic integration site preference of retroviral vectors has been the major bottleneck as it was shown to be the cause of vector induced adverse events in early clinical trials. For these reasons I embarked on two parallel research lines with a main focus on the intricate interplay between LEDGF/p75-IN, contributing to the development of safer viral vectors for gene therapy and alternative HIV cure strategies that modulate or prevent the generation of a reservoir able to rebound from latency once drug administration is withdrawn.

I) Development of Safer LV Vectors Using Artificial LEDGF-Based Tethers

Lentiviral integration occurs in the body of active transcription units [Schroder et al., 2002; Mitchell et al., 2004]. Previously, we and others showed that integration can be retargeted towards regions generally disfavored for integration by using artificial LEDGF-based tethers [Ferris et al., 2010; Gijsbers et al., 2009, 2011b; Silvers et al., 2010]. In a first part of my work (Chapter 3 on page 41), I set out to design new chimeric versions of LEDGF/p75 to retarget integration to distinct nuclear compartments/locations and potentially 'safer' regions within the human genome. As such, serious adverse events associated with the use of retroviral vectors in gene therapeutic applications may be overcome [Hacein-Bey-Abina et al., 2003b, 2008; Howe et al., 2008; Stein et al., 2010]. In light of a safer genotoxicity profile, I tested whether truncation of LEDGF/p75 by deletion the N-terminal chromatin-reading PWWP domain (responsible for binding H3K36me3 marks associated with active transcription units) could result in a more randomized integration profile. A random integration is generally believed to reduce the risk or probability of transcriptional dysregulation of neighbouring genes. In parallel, I replaced the PWWP-domain with several alternative viral protein domains and motifs, described in literature as pan-chromatin recognition peptides since they bind cellular chromatin without sequence specificity aiding episomal viruses to persist during mitosis; The spumavirus, Prototype Foamy Virus (PFV), contains a 13-amino acid motif in the group-specific antigen (Gag) binding the H2A/H2B core nucleosome [Tobaly-Tapiero et al., 2008; Nowrouzi et al., 2006; Trobridge et al., 2006]. Likewise, the Kaposi Sarcoma-associated Herpes Virus (KSHV) genome is tethered to the nucleosomal core via a chromatin binding sequence (CBS) at the N-terminal end of the latency-associated nuclear antigen protein (LANA) [Barbera et al., 2006]. Finally, in the β-Papillomaviruses (PV) a conserved motif in

the E2 hinge promotes binding to chromatin and mitotic chromosomes of the invaded cell [Sekhar et al., 2010; Sekhar and McBride, 2012; Võsa et al., 2012]. Retargeting was confirmed by proviral integration site sequencing.

II) Study of The Impact of Integration Site Distribution on HIV Latency

The fact that lentiviral integration can be retargeted to genomic regions that are usually disfavoured for integration indicates that these areas are disfavoured in WT cells due to active LEDGF/p75 tethering rather than an inherent integration barrier such as steric hindrance resulting from condensed chromatin. In a second research line (Chapter 4 on page 69), I studied the importance of integration site selection/distribution in the context of HIV persistence, evaluating (I) the establishment of latency and (II) the reactivation from latency following treatment with Latency Reversing Agents (LRAs). Since integration is believed to affect the transcriptional activity of the integrated provirus, I assessed whether targeting and integration site selection affect the reactivation potential of the quiescent integrated provirus and explored the relationship with chromatin structure and nuclear topography by modulating the LEDGF/p75-IN interplay. Moreover, I investigated the hypothesis that LEDGIN treatment, small molecules abrogating the LEDGF/p75 interaction, could result in re-targeted integration and studied their effect on the formation of a functional HIV reservoir. A better characterization of their mechanisms of action might aid their further clinical development and contribute to the development of alternative strategies modulating or preventing the generation of latent reservoirs able to rebound viremia and thereby help to overcome the HIV pandemic.

3

Towards a Safer, More Randomized Lentiviral Integration Profile Exploring Artificial LEDGF Chimeras

This chapter has been previously published as an article manuscript in PLoS One (2016):

Vranckx, L.S., Demeulemeester, J., Gijsbers, R., Debyser, Z. (2016). Towards a Safer, More Randomized Lentiviral Vector Integration Profile Exploring Artificial LEDGF Chimeras. PLoS One, 11(10), art.nr. 10.1371/journal.pone.0164167.

3.1 Abstract

The capacity to integrate transgenes into the host cell genome makes retroviral vectors an interesting tool for gene therapy. Although stable insertion resulted in successful correction of several monogenic disorders, it also accounts for insertional mutagenesis, a major setback in otherwise successful clinical gene therapy trials due to leukemia development in a subset of treated patients. Despite improvements in vector design, their use is still not risk-free. Lentiviral vector (LV) integration is directed into active transcription units by LEDGF/p75, a host-cell protein co-opted by the viral integrase. We engineered LEDGF/p75-based hybrid tethers in an effort to elicit a more random integration pattern to increase biosafety, and potentially reduce of proto-oncogene activation. We therefore truncated LEDGF/p75 by deleting the N-terminal chromatin-reading PWWP-domain, and replaced this domain with alternative pan-chromatin binding peptides. Expression of these LEDGF-hybrids in LEDGF-depleted cells efficiently rescued LV transduction and resulted in LV integrations that distributed more randomly throughout the host-cell genome. In addition, when considering safe harbor criteria, LV integration sites for these LEDGF-hybrids distributed more safely compared to LEDGF/p75-mediated integration in wild-type cells. This approach should be broadly applicable to introduce therapeutic or suicide genes for cell therapy, such as patient-specific iPS cells.

3.2 Introduction

The capacity to integrate transgenes into the host cell genome makes Retroviral Vectors (RV) an interesting tool for gene therapeutic applications as stable insertion of transgenes into the genome ensures long-term expression. Use of RV-mediated gene transfer resulted in successful cure of several monogenic, primary immunodeficiency disorders [Cavazzana-Calvo et al., 2000; Gaspar et al., 2004; Hacein-Bey-Abina et al., 2003b]. Yet, stable insertion occasionally altered endogenous gene regulation resulting in insertional mutagenesis. Due to this major setback 25% of treated patients developed leukemia in otherwise successful clinical gene therapy trials for X-SCID and X-CGD [Hacein-Bey-Abina et al., 2003b, 2008; Howe et al., 2008; Stein et al., 2010]. Both trials employed Murine Leukemia Virus (MLV)-based gammaretroviral vectors (γRV) that integrate in close proximity to gene regulatory regions [Wu et al., 2003; De Ravin et al., 2014; LaFave et al., 2014] and resulted in transcriptional deregulation due to up-regulated LMO2 expression [Mitchell et al., 2004; Derse et al., 2007; Deichmann et al., 2007; Cattoglio et al., 2007]. Similar reports on insertional mutagenesis were published after integration of RV near CCDN2, BMI1 and EVI1 [Boztug et al., 2010; Braun et al., 2014]. Despite improvements in vector design (e.g. self-inactivating (SIN) vectors) their use is still not risk-free [Hacein-Bey-Abina et al., 2003b, 2008; Stein et al., 2010; Boztug et al., 2010; Braun et al., 2014; Maetzig et al., 2011], which shifted attention from γRV towards HIV-derived lentiviral vectors (LV). Even though LV display a more favorable integration pattern, induction of aberrant splicing [Cesana et al., 2012, 2014] and insertional mutagenesis remain a major concern, as clonal expansion was observed in a gene therapy trial for β-thalassemia [Cavazzana-Calvo et al., 2010]. In addition, two recent independent studies revealed

clonal expansion in HIV-1 infected patients on antiretroviral therapy due to HIV-1 virus triggered insertional mutagenesis [Maldarelli et al., 2014; Wagner et al., 2014]. Retroviral integration is a non-random process which is, depending on the viral genus, associated with specific chromatin marks and genomic features [Stevens and Griffith, 1996; Holman and Coffin, 2005; Wu et al., 2005]. γRV predominantly integrate in the vicinity of gene regulatory regions, whereas LV preferably target the body of active transcription units [Schroder et al., 2002; Mitchell et al., 2004]. Integration is catalyzed by the viral INtegrase (IN), whereas integration site choice bias is attributed to the cellular chromatin readers that are co-opted by the viral IN. Whereas the Bromodomain and Extra-Terminal domain (BET) family of proteins (BRD2, 3 and 4) guide MLV integration [Sharma et al., 2013; De Rijck et al., 2013; Gupta et al., 2013], LV integration is directed by Lens Epithelium-Derived Growth Factor p75 (LEDGF/p75) [Cherepanov et al., 2003; Marshall et al., 2007]. Both function as molecular tethers in the cell, combining a chromatin-binding and a protein-interacting region (reviewed in [Debyser et al., 2015]). For LEDGF/p75 (Figure 3.1a on page 44), the chromatin-binding part contains an N-terminal Pro-Trp-Trp-Pro (PWWP) epigenetic reader domain (aa 1-93), recognizing H3K36me3 chromatin marks [De Rijck et al., 2010; Pradeepa et al., 2012; Eidahl et al., 2013; Gijsbers et al., 2011a; van Nuland et al., 2013], and a set of DNA-binding motifs (Figure 3.1a, [Turlure et al., 2006; Tsutsui et al., 2011]). Together, these elements allow LEDGF/p75 to explore the chromatin in a dynamic scan-and-lock fashion [Hendrix et al., 2011]. Even though its cellular role is not fully understood, it is clear that LEDGF/p75 acts as a molecular hub for a variety of endogenous proteins next to the lentiviral integrase (Figure 3.1a) [Maertens et al., 2006; Bartholomeeusen et al., 2007, 2009; Hughes et al., 2010; Yokoyama and Cleary, 2008]. All these proteins, including the lentiviral integrase, bind the C-terminal Integrase-Binding Domain (IBD, aa 347-429; Figure 3.1a) of LEDGF/p75. We and others showed that replacement of the N-terminal LEDGF/p75 DNA-binding region (aa 1-325) with alternative DNA-binding domains retargets LV integration towards genomic loci bound by these domains [Ferris et al., 2010; Gijsbers et al., 2009, 2011b; Silvers et al., 2010]. Fusion of the heterochromatin binding Chromobox protein homolog 1 (CBX1) to the IN-binding C-terminal end of LEDGF/p75 shifted LV integration into the cognate H3K9me$_x$-marked chromatin environment, pericentric heterochromatin and intergenic regions [Gijsbers et al., 2009]. Despite integration in regions enriched in epigenetic marks associated with gene silencing, transgene expression remained efficient and resulted in successful phenotypic correction in a cell model for X-CGD [Vets et al., 2013]. Here we aimed at developing a LEDGF-based tether that results in a more random integration pattern to reduce the overall risk of insertional mutagenesis [Chatziandreou et al., 2011; Derse et al., 2007; Kaufmann et al., 2013; Suerth et al., 2012]. First, we truncated LEDGF/p75 by deleting the N-terminal chromatin-reading PWWP domain that binds H3K36me3 marks directing LEDGF/p75 into the body of active transcription units (Figure 3.1a & b). In addition, we replaced the PWWP-domain with three alternative viral protein domains and motifs, described in literature as pan-chromatin recognition peptides since they bind cellular chromatin without sequence specificity (Figure 3.1b and Table 3.1 on page 44 and 45 respectively). Several viruses reside as an episomal DNA genome in host cells, and evolved strategies to persist during mitosis through defined chromatin binding motifs. The spumavirus, Prototype Foamy Virus (PFV), contains a 13-amino acid

motif in the group-specific antigen (Gag) binding the H2A/H2B core nucleosome [Tobaly-Tapiero et al., 2008; Nowrouzi et al., 2006; Trobridge et al., 2006]. Likewise, the Kaposi Sarcoma-associated Herpes Virus (KSHV) genome is tethered to the nucleosomal core via a Chromatin Binding Sequence (CBS) at the N-terminal end of the Latency-Associated Nuclear Antigen protein (LANA) [Barbera et al., 2006]. Finally, in the β-PapillomaViruses (PV) a conserved motif in the E2 hinge promotes binding to chromatin and mitotic chromosomes of the invaded cell [Sekhar et al., 2010; Sekhar and McBride, 2012; Võsa et al., 2012]. Following the generation of stable cells lines, we monitored LV integration preferences and evaluated integration sites based on safe harbor region criteria [Papapetrou et al., 2011] and determined a genotoxicity profile.

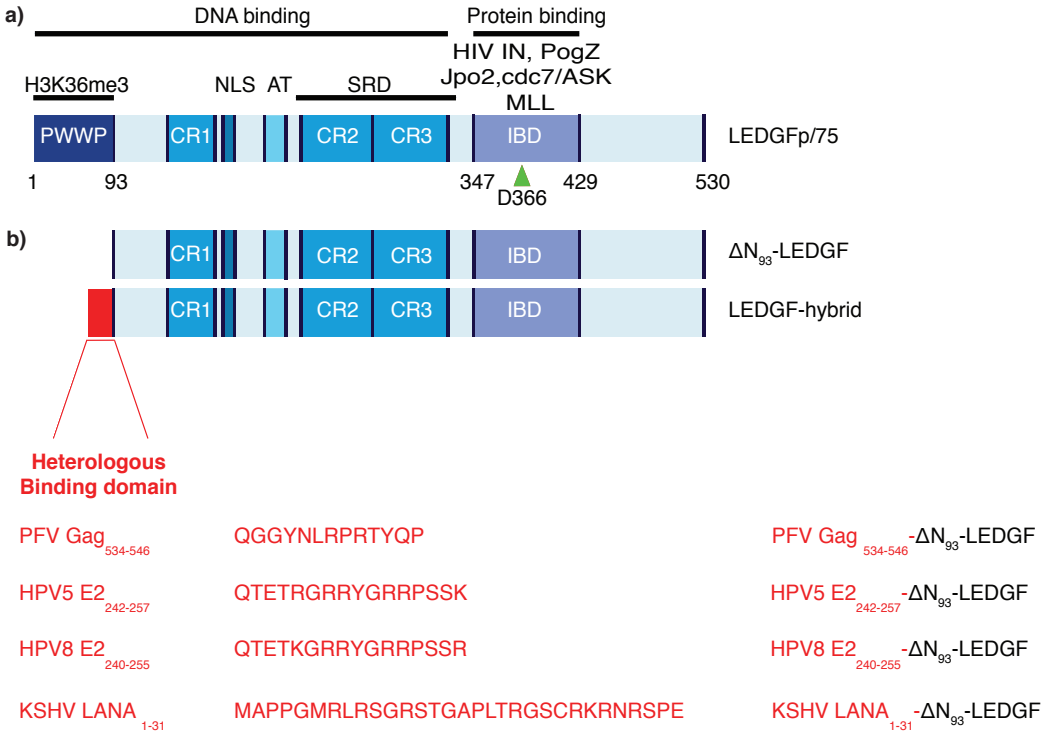


Figure 3.1: Schematic representation of the LEDGF/p75 domain structure and artificial LEDGF-hybrids. (a) LEDGF/p75 contains a C-terminal protein-binding domain, coined Integrase Binding Domain (IBD) responsible for HIV-IN interaction. Several endogenous proteins like Jpo2, PogZ and MLL bind to the same interface. At its N-terminal end carries multiple chromatin interacting domains, the PWWP domain, the AT hook-like domain (AT) and three Charged Regions (CR1, 2, 3). D366 is a pivotal amino acid involved in HIV-IN interaction (green arrowhead). Mutation to Asn (D366N) abolishes HIV-IN interaction. The lower panel (b) depicts the different LEDGF-hybrids, PFV $Gag_{534\text{-}546}$-δN_{93}-LEDGF, HPV5 $E2_{242\text{-}257}$-δN_{93}-LEDGF, HPV8 $E2_{240\text{-}255}$-δN_{93}-LEDGF and $LANA_{1\text{-}31}$-δN_{93}-LEDGF respectively. Numbers indicate the different amino acid residues. AT, AT-Hook; CR, Charged Region; SRD, Supercoiled Recognition Domein; IBD, Integrase Binding Domain; PWWP, Pro-Trp-Trp-Pro Domain; PFV, Prototype Foamy Virus; LANA, Latency Associated Nuclear Antigen; HPV, Human Papilloma Virus; LEDGF, Lens Epithelium-Derived Growth Factor; NLS, Nuclear Localization Signal.

Peptide	Name	Sequence	Target	Reference
PFV $Gag_{534-546}$	Prototype foamy virus chromatin binding segment Gag	QGGYNLRPRTYQP	H2A-H2B core histones	Tobaly-Tapiero 2008, Nowrouzi 2006, Trobridge 2006
HPV5 $E2_{242-257}$	Human papilloma virus E2 protein	QTET**R**GRRYGRRPSS**K**	12-bp DNA motifs-E2 binding sites (E2BS) $ACCN_6GGT$	Sekhar and McBride, 2012; Sekhar et al., 2010; Vosa 2012
HPV8 $E2_{240-255}$	Human papilloma virus E2 protein	QTET**K**GRRYGRRPSS**R**	12-bp DNA motifs-E2 binding sites (E2BS) $ACCN_6GGT$	Sekhar and McBride, 2012; Sekhar et al., 2010; Vosa 2012
KSHV $LANA_{1-31}$	Kaposi sarcoma herpes virus latency associated nuclear antigen	MAPPGMRLRSGRSTGAPLTRGSCRKR NRSPE	H2A-H2B nucleosome core	Barbera Science 2006

Table 3.1: Peptide characteristics. Table showing the acronyms, aa-sequences and binding characteristics of the peptides used to generate the artificial LEDGF-tethers. PFV, Prototype Foamy Virus; HPV, Human Papilloma Virus; KSHV, Kaposi's Sarcoma Herpes Virus; LANA, Latency Associated Nuclear Antigen.

3.3 Results

3.3.1 Generation of LEDGF-hybrids and stable cell lines

In an effort to distribute lentiviral vector integration more randomly over the genome, we modified LEDGF/p75, the cellular tether of the HIV Pre-Integration Complex (PIC), by deleting the chromatin-reading PWWP-domain (δN_{93}-LEDGF, Figure 3.1a, b on page 44) [De Rijck et al., 2010; Pradeepa et al., 2012; Eidahl et al., 2013] relying on the remaining non-specific DNA-interacting regions in $LEDGF_{93-325}$, such as the AT-hook domains and the CRs (Figure 3.1a; [Hendrix et al., 2011]). In addition, we generated LEDGF-hybrids where the PWWP-domain was exchanged with a set of alternative pan-chromatin recognition peptides of viral origin (Figure 3.1b and Table 3.1 on page 45). Prototype Foamy Virus chromatin binding segment of $Gag_{534-546}$ (PFV $Gag_{534-546}$) [Tobaly-Tapiero et al., 2008; Nowrouzi et al., 2006; Trobridge et al., 2006], Human Papilloma Virus serotype 5 $E2_{242-257}$, Human Papilloma Virus serotype 8 $E2_{240-255}$ (HPV5 $E2_{242-257}$ and HPV8 $E2_{240-255}$, respectively)[Sekhar et al., 2010; Sekhar and McBride, 2012; Võsa et al., 2012] and Kaposi's Sarcoma Herpes Virus Latency Associated Nuclear $Antigen_{1-31}$ (KSHV $LANA_{1-31}$) [Barbera et al., 2006] were used to replace the PWWP domain, generating δN_{93}-LEDGF, PFV $Gag_{534-546}$-δN_{93}-LEDGF, HPV5 $E2_{242-257}$-δN_{93}-LEDGF, HPV8 $E2_{240-255}$-δN_{93}-LEDGF and KSHV $LANA_{1-31}$-δN_{93}-LEDGF fusions, respectively. All above-mentioned LEDGF-hybrids were used to complement LEDGF/p75-depleted cells (HeLaP4 ($LEDGF_{KD}$) [Osório et al., 2014] and Nalm ($LEDGF_{KO}$) cells [Schrijvers et al., 2012a]) employing SIV-based lentiviral vectors. As a positive control, cells were complemented with WT LEDGF/p75 (referred to as LEDGF/p75 back complementation ($LEDGF_{BC}$)). In order to control for non-specific effects resulting from the expression of the fusion proteins we also generated stable cell lines expressing the respective chimeras carrying a D366N mutation in the LEDGF/p75 part, which abrogates the interaction with lentiviral integrase (IN) [Cherepanov et al., 2005]. Protein integrity was corroborated by Western Blot (WB) analysis, with all LEDGF-hybrids migrating at the predicted molecular weights (Supplementary Figure S3.1 on page 63). Of note, protein levels of PFV $Gag_{534-546}$-δN_{93}-LEDGF were lower in all experiments. Viability and growth rates of all cell lines were comparable to the parental HeLaP4 cells (data not shown).

3.3.2 LEDGF-hybrids locate to the nucleus and display a distinct subnuclear distribution

In a first step, we evaluated the subcellular localization of the truncated δN_{93}-LEDGF and the respective δN_{93}-LEDGF-hybrids by immunocytochemistry (Figure 3.2 on page 47). Complementation of LEDGF-depleted HeLaP4 cells ($LEDGF_{KD}$) with $LEDGF_{BC}$ resulted in a typical pattern of dense, fine speckles in the nucleoplasm excluded from the nucleoli during interphase (Figure 3.2c), phenocopying the endogenous LEDGF/p75 pattern (Figure 3.2a), which is in line with earlier reports [Gijsbers et al., 2009]. On the contrary, LEDGF/p75 lacking the chromatin-reading PWWP-domain exhibited a more diffuse nuclear distribution and located to the nucleoli as well (δN_{93}-LEDGF, Figure 3.2d). In addition, all δN_{93}-LEDGF peptide-fusions located to the nucleus (Figure 3.2e-h), displaying a unique sub-nuclear distribution: the PFV $Gag_{534\text{-}546}$- and the KSHV $LANA_{1\text{-}31}$-fusion to δN_{93}-LEDGF showed a punctate appearance in the nucleus and were excluded from nucleoli (Figure 3.2e, h), contrary to both HPV5 $E2_{242\text{-}257}$- and HPV8 $E2_{240\text{-}255}$-δN_{93}-LEDGF fusions that were enriched in the nucleoli (Figure 3.2f, g). Similar subcellular distributions were observed for the respective cognate $LEDGF_{D366N}$-hybrids (data not shown).

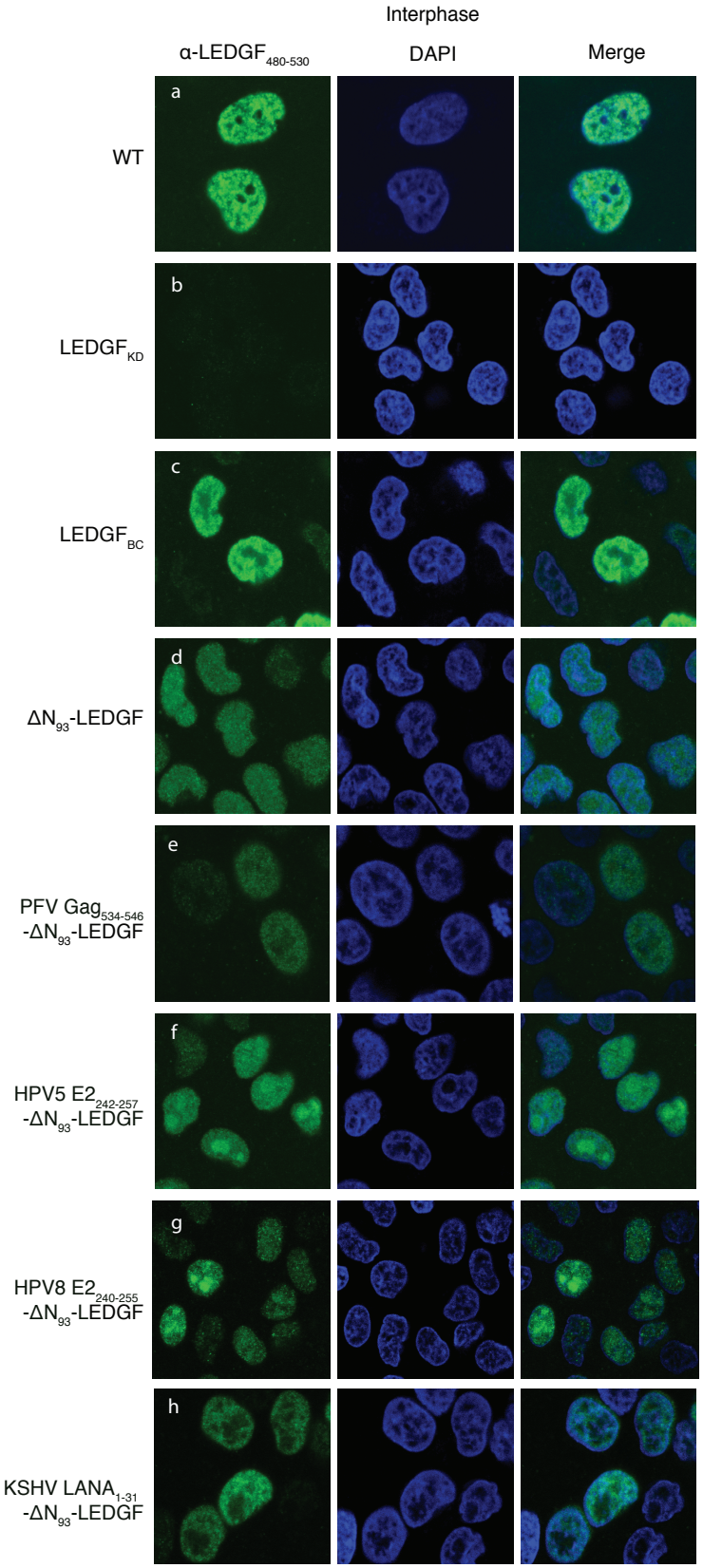


Figure 3.2: Subcellular localization of LEDGF-hybrids in interphase cells. LEDGF/p75 depleted cell lines were complemented with the respective LEDGF-fusions. Laser scanning confocal images of HelaP4 cells, stained using an Ab recognizing $LEDGF_{480\text{-}530}$, are shown in green. Nuclei were stained using DAPI (shown in blue). A merge of green and blue fluorescence is shown. Data depicted are representative for the respective cell lines. PFV, Prototype Foamy Virus; LANA, Latency Associated Nuclear Antigen; HPV, Human Papilloma Virus; LEDGF, Lens Epithelium-Derived Growth Factor; DAPI, 4', 6-DiAmidino-2-PhenylIndole.

3.3.3 LEDGF-peptide fusions rescue lentiviral vector transduction

Next, we assessed whether the δN_{93}-LEDGF-hybrids supported lentiviral vector transduction by complementing LEDGF-depleted cells ($LEDGF_{KD}$) and employing wild-type LEDGF/p75 complemented cells ($LEDGF_{BC}$) as control. The respective HeLaP4 cell lines were challenged with a dilution series of a lentiviral vector (multiplicity of infection (MOI)= 1, 0.2 or 0.04 (indicated in lighter colors)) encoding enhanced Green Fluorescent Protein (eGFP) and firefly Luciferase (fLuc) reporters [Ibrahimi et al., 2009]. Transduction efficiencies were determined by flow cytometry monitoring eGFP fluorescence (Figure 3.3a and b on page 49, showing transduction efficiency (eGFP positive cells; % Gated) and Mean Fluorescence Intensity (MFI) respectively). Complementation of LEDGF-depleted cells with $LEDGF_{BC}$ restored transduction efficiency (Figure 3.3a)(***, $p<0.005$; two-tailed t-test relative to $LEDGF_{KD}$), in line with earlier reports [Vandekerckhove et al., 2006; Gijsbers et al., 2009]. Complementation of $LEDGF_{KD}$ cells with δN_{93}-LEDGF, lacking the chromatin interacting PWWP-domain, partially rescued lentiviral transduction (78% compared to $LEDGF_{BC}$) (Figure 3.3a, ***, $p<0.005$ compared to KD, two-tailed t-test). Addition of chromatin binding peptides to replace the PWWP domain displayed a significantly improved transduction relative to δN_{93}-LEDGF (***, $p<0.005$; two-tailed t-test), reaching efficiencies comparable to $LEDGF_{BC}$ (Figure 3.3a). Similar results were obtained for different vector dilutions (Figure 3.3 a) or when assessing complemented LEDGF/p75 knock-out cells ($Nalm^{-/-}$, data not shown) [Schrijvers et al., 2012a] or evaluating fLuc as a reporter (data not shown). Looking at Mean Fluorescence Intensities (MFI), all LEDGF-peptide fusions were about 20 % lower than $LEDGF_{BC}$ (Figure 3.3b). In addition to transduction efficiencies, we also determined the number of integrated copies (Figure 3.3c). Reintroduction of LEDGF/p75 ($LEDGF_{BC}$) significantly improved vector integration ($\pm$3.5-fold compared to $LEDGF_{KD}$, ***, $p<0.005$, two-tailed t-test). Likewise, δN_{93}-LEDGF and all LEDGF-peptide fusions restored vector integration (2.5-fold more compared to $LEDGF_{KD}$, ***, $p<0.005$, two-tailed t-test), albeit still to a lesser extent (reaching 68-74% of $LEDGF_{BC}$, Figure 3.3c). The increased transduction efficiencies (%Gated) closely correlate with an increase in integrated viral vector copies (Figure 3.3c). Complementation with δN_{93}-LEDGF alone, lacking any additional chromatin-tether, resulted in lower integrated copy numbers than δN_{93}-LEDGF fused to chromatin engaging peptides (p-values < 0.005, two-tailed t-test relative to δN_{93}-LEDGF; 66.7% of $LEDGF_{BC}$), supporting the notion that the chromatin-reading PWWP domain of LEDGF/p75 is not an absolute requirement for efficient integration.

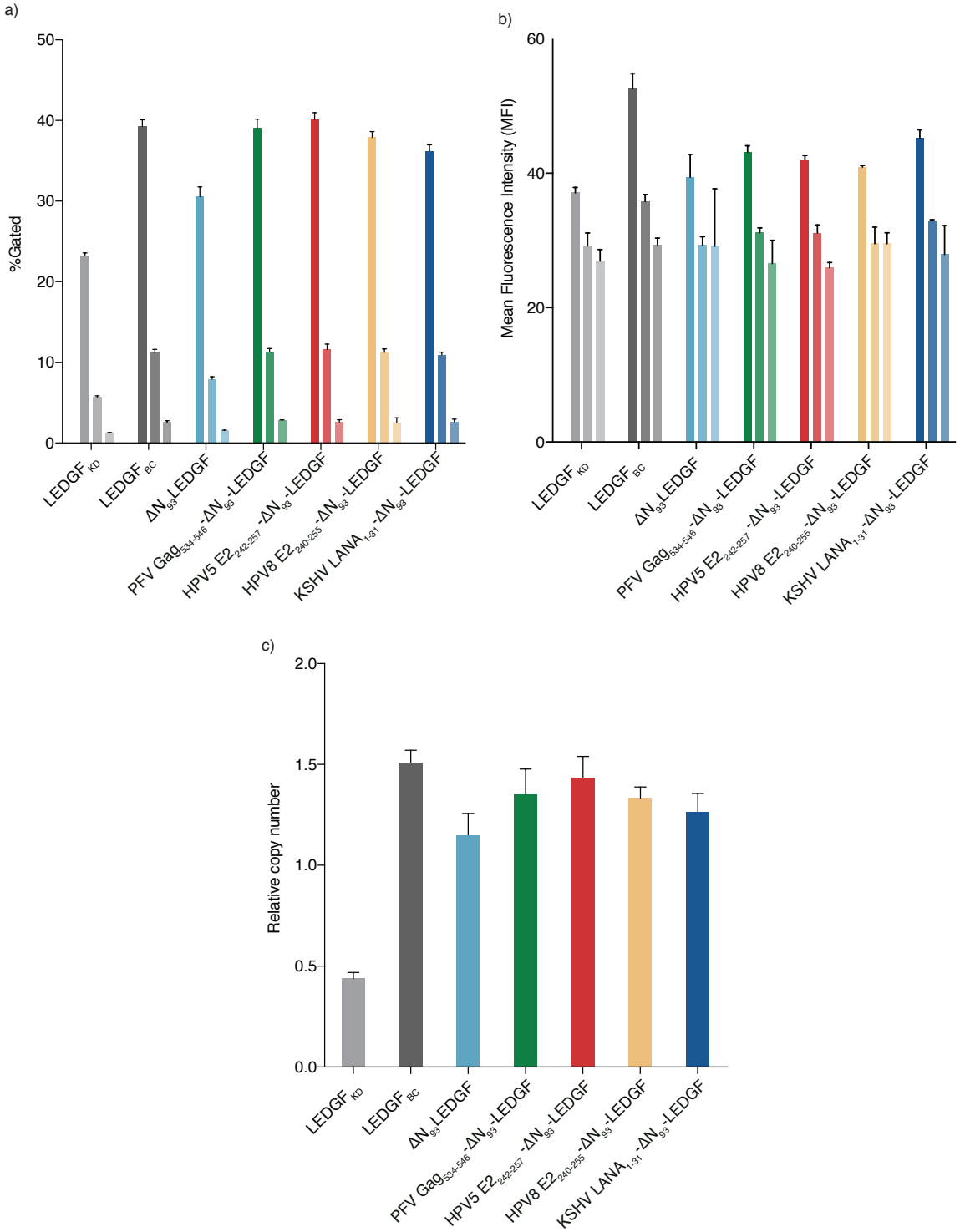


Figure 3.3: Rescue of lentiviral vector transduction by artificial LEDGF-hybrids. LEDGF-fusions were evaluated for their ability to support lentiviral vector transduction. LEDGF-depleted HelaP4-based cell lines stably complemented with LEDGF-hybrids were challenged with a VSV-G pseudo-typed lentiviral reporter vector encoding enhanced Green Fluorescent Protein (eGFP). Fluorescence was measured by fluorescence activated cell sorting and the different variables plotted: (a) Percentage eGFP positive cells (transduction efficiency) and (b) Mean Fluorescence Intensity (MFI). Data are compiled for a representative experiment and depict averages of 3 replicates for 3 different vector dilutions (mean ± SD). (c) Lentiviral integrated proviral copies were determined by Q-PCR analysis on genomic DNA extracts of cells transduced with an MOI=1. Data are representeded as the mean of 3 replicates ± SD. Statistical significance is calculated using a two-tailed t-test relative to $LEDGF_{KD}$ or δN_{93}-LEDGF. PFV, Prototype Foamy Virus; LANA, Latency Associated Nuclear Antigen; HPV, Human Papilloma Virus; LEDGF, Lens Epithelium-Derived Growth Factor.

3.3.4 LEDGF-peptide fusions efficiently redistribute lentiviral integration

After showing that complementation of LEDGF/p75-depleted cells with δN_{93}-LEDGF or any of the LEDGF-hybrids rescued vector integration, we determined the integration profiles in the respective cell lines. HIV-based viral vector integration sites were amplified and sequenced as described earlier [Marshall et al., 2007; Gijsbers et al., 2009], yielding a total of 62670 unique integration sites and their computationally generated Matched Random Control (MRC) sites. Note that SIV-based viral vectors were used to complement LEDGF/p75-depleted cells, in order to avoid interference with the HIV-based viral vector integration site amplification and analysis. First, we analysed integration relative to a set of defined genomic features (Table 3.2 and Figure 3.4 on page 52 and 53 respectively). Lentiviral vector integration in wild-type HeLaP4 cells (endogenous LEDGF/p75, Figure 3.4) is traditionally enriched in the body of transcription units (75.0% in RefSeq genes; Table 3.2) but disfavoured Transcription Start Sites (TSS) and promoter regions (2.0% within 2kb of the 5' of a RefSeq gene and 3.1% within 2kb of a CpG island) [Schroder et al., 2002; Mitchell et al., 2004]. LEDGF-depletion results in a more random integration site distribution, characterized by reduced integration into genes (51.0% in RefSeq genes) and increased integration close to TSS (5.4%) and CpG islands (7.0%), in line with previous work [Shun et al., 2007; Marshall et al., 2007; Gijsbers et al., 2009]. This phenotype was fully reverted upon LEDGF/p75 complementation ($LEDGF_{BC}$; 75.6% in RefSeq genes). Comparable data were obtained for larger window sizes (only 2kb and 4kb are shown in Table 3.2). Integration site distributions in cells expressing the respective $LEDGF_{D366N}$ mutants were not different from $LEDGF_{KD}$ cells (n=16473; data not shown). Interestingly, the mere ablation of the PWWP domain (δN_{93}-LEDGF) resulted in an overall more random distribution compared to $LEDGF_{KD}$ cells, with decreased integration near retrovirus-specific features like gene bodies, TSS and promoter regions (*** $p<0.001$; χ^2 test compared to $LEDGF_{KD}$; Table 3.2). Complementation of LEDGF-depleted cells with LEDGF-peptide fusions resulted in a comparable more randomized distribution (*** $p<0.001$; χ^2 test compared to $LEDGF_{KD}$; Table 3.2). In a more elaborate analysis, we analysed global integration preferences and included a wide selection of genomic features, depicted as a genomic heatmap (Figure 3.4), comparing integration site data sets obtained from HeLaP4 $LEDGF_{KD}$ cells to those of cells complemented with the respective LEDGF-hybrids. Tile color depicts the correlation for an integration dataset with the respective genomic feature (left) relative to matched random controls, as indicated by the colored Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. LEDGF/p75 depletion shifts integration out of transcriptionally active regions wich is reverted upon complementation with LEDGF/p75 (compare $LEDGF_{KD}$ and $LEDGF_{BC}$; shown in Figure 3.4 and Supplementary Figure S3.2 on page 64 (compared to $LEDGF_{BC}$ and δN_{93}-LEDGF respectively), in line with previous data [Shun et al., 2007; Marshall et al., 2007; Gijsbers et al., 2009]. Cells complemented with δN_{93}-LEDGF displayed an more randomly distributed integration profile, with tiles overall coloring less red or blue compared to $LEDGF_{KD}$, integrating less near DNase sensitive regions, CpG-islands and GC-rich regions compared to $LEDGF_{KD}$ (*** $p<0.001$, Wald statistics). Introduction of the heterologous HPV E2 and $LANA_{1\text{-}31}$-peptide fragments to replace the PWWP-domain resulted in a δN_{93}-LEDGF-like integration

profile when compared to $LEDGF_{KD}$ ($p<0.001$), whereas integration for PFV $Gag_{534\text{-}546}$-δN_{93}-LEDGF was less random. When displaying statistics relative to δN_{93}-LEDGF (Supplementary Figure S3.2a) it is clear that integration frequencies relative to these genomic features is not significantly different between δN_{93}-LEDGF and the respective δN_{93}-LEDGF peptide-fusions, except for PFV $Gag_{534\text{-}546}$-δN_{93}-LEDGF, which shifts to more random relative to $LEDGF_{KD}$. The reproducibility of the data observed for HPV5 $E2_{242\text{-}257}$-δN_{93}-LEDGF and HPV8 $E2_{240\text{-}255}$-δN_{93}-LEDGF complemented cell lines and the pronounced redistribution towards more random relative to $LEDGF_{BC}$ (Supplementary Figure S3.2b) underscores the effectiveness of LEDGF-based artificial tethers for retargeting of LV integration. Next to integration relative to genomic features, we also analyzed integration site densities near epigenetic features (Figure 3.4b on page 54). The epigenetic heat map displays yellow and blue tiles, with blue tiles indicating that integration frequency is enriched near these marks relative to MRC, whereas yellow tiles indicate that integration is disfavored compared to MRC. A near random distribution would result in a black tile. As reported previously, lentiviral integration correlates with histone marks associated with open and transcriptionally active chromatin (H3K4 mono-, di- and tri methylation, H3K14 and H4 acetylation, as well as acetylation and monomethylation of H3K9/K27/K79, H4K20 and H2BK5,...)[De Ravin et al., 2014] while disfavoring integration in transcriptionally silent regions or heterochromatin ($H3K27me_3$, $H3K9me_3$ or $H4K20me_3$ and H3K79, respectively) [De Ravin et al., 2014] (WT; Figure 3.4b). Depletion of LEDGF/p75 ($LEDGF_{KD}$) resulted in a more random distribution (with tiles displaying a less pronounced blue or yellow color, and shifting towards black). This tendency was more outspoken for δN_{93}-LEDGF and the HPV E2 and $LANA_{1\text{-}31}$-peptide fusions compared to $LEDGF_{KD}$ (Figure 3.4, δN_{93}-LEDGF (Supplementary Figure S3.3a) or $LEDGF_{BC}$ (Supplementary Figure S3.3b on page 65), potentially because integration in LEDGF-depleted cells, at least in part, is tethered by HRP-2 [Schrijvers et al., 2012a].

	Type	Controls	Total Sites	InRefGene	% TSS within 2kb	% TSS within 4Kb	% CpG within 2Kb	% CpG within 4Kb	% DHS within 2Kb	% DHS within 4Kb
WT	insertion	FALSE	3520	75.0 ***	2.0 ***	7.4 ***	3.1 ***	9.5 ***	23.0	40.7
$LEDGF_{KD}$	insertion	TRUE	5560	51.0	5.4	11.4	7.0	12.3	22.4	37.8
$LEDGF_{BC}$	insertion	FALSE	1107	75.6 ***	2.0 ***	6.5 ***	2.7 ***	8.2	18.3	35.3
ΔN_{93}-LEDGF	insertion	FALSE	2990	51.0	2.3 ***	6.3 ***	2.7 ***	6.5 ***	16.9 ***	31.8 ***
PFV $Gag_{534-546}$-ΔN_{93}-LEDGF	insertion	FALSE	8267	51.1	2.7 ***	6.7 ***	3.4 ***	7.4 ***	18.5 ***	33.4 ***
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	insertion	FALSE	9572	48.6	2.2 ***	5.5 ***	2.2 ***	5.7 ***	16.2 ***	29.8 ***
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	insertion	FALSE	6616	48.1	2.0 ***	5.4 ***	2.3 ***	5.5 ***	16.7 ***	29.8 ***
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	insertion	FALSE	8565	48.4	1.4 ***	4.6 ***	2.2 ***	5.2 ***	16.4 ***	29.6 ***
WT	match	FALSE	10553	39.5	2.3	4.8	2.9	5.1	15.2	27.0
$LEDGF_{KD}$	match	TRUE	16680	39.7	2.2	4.4	2.6	5.0	14.4	26.2
$LEDGF_{BC}$	match	FALSE	3321	40.1	2.2	4.9	2.7	5.3	15.0	26.3
ΔN_{93}-LEDGF	match	FALSE	8970	39.7	2.6	5.2	3.1	5.9	15.4	27.2
PFV $Gag_{534-546}$-ΔN_{93}-LEDGF	match	FALSE	24792	39.7	2.2	4.7	2.9	5.2	14.9	26.6
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	match	FALSE	28716	40.0	2.1	4.6	2.6	5.1	15.1	27.0
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	match	FALSE	19845	40.1	2.1	4.7	2.8	5.4	15.0	26.4
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	match	FALSE	25690	39.9	2.2	4.4	2.7	5.0	15.0	26.6

Table 3.2: Integration frequency near genomic features. Table showing the percentage HIV-derived lentiviral vector integration sites relative to features specific for integrating viral vectors such as integration into the body of genes (Refseq genes, InRefGene), integration within 2kb-4kb windows near Transcription Start Sites (X5-end of genes,TSS), midpoint of CpG islands or DNase I-hypersensitive sites (DHS), counted in both the 5' as 3' direction. Dataset details are described in the MM section. Asterisks depict a significant deviation from $LEDGF_{KD}$ (two-tailed Chi-square test; ***, p-values <0.001). TSS, Transcription Start Sites; DHS, DNase I-Hypersensitive Sites; PFV, Prototype Foamy Virus; HPV, Human Papilloma Virus; KSHV, Kaposi's Sarcoma Herpes Virus; LANA, Latency Associated Nuclear Antigen; match, matched random control.

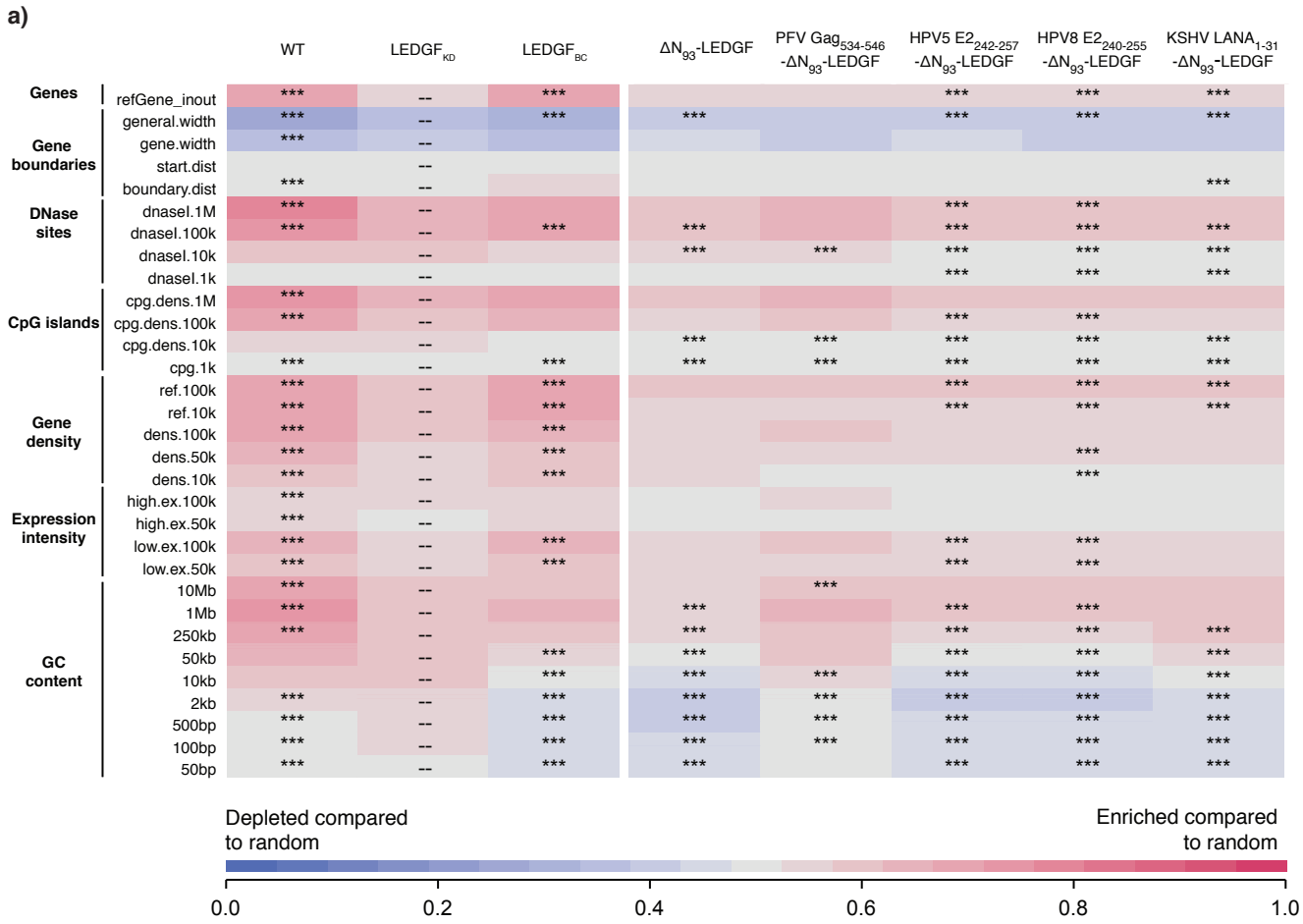


Figure 3.4: LEDGF-hybrids retarget lentiviral integration towards a more randomized pattern. (a) Genomic heat map comparing integration site data sets obtained from HeLaP4 LEDGF/p75 KD cells overexpressing different artificial LEDGF-hybrids to genomic features. Tile color depicts the correlation for an integration dataset with the respective genomic feature (left) relative to matched random controls, as indicated by the colored Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. Statistical significance (asterisks, ***p<0.001 ranked Wald tests) is shown relative to $LEDGF_{KD}$ population (double dash). Columns indicate different data sets, while rows indicate different genomic features analyzed (described in [Ocwieja et al., 2011]). LANA, Latency Associated Nuclear antigen; HPV, Human Papilloma Virus; PFV, Prototype Foamy Virus; LEDGF, Lens Epithelium-Derived Growth Factor;

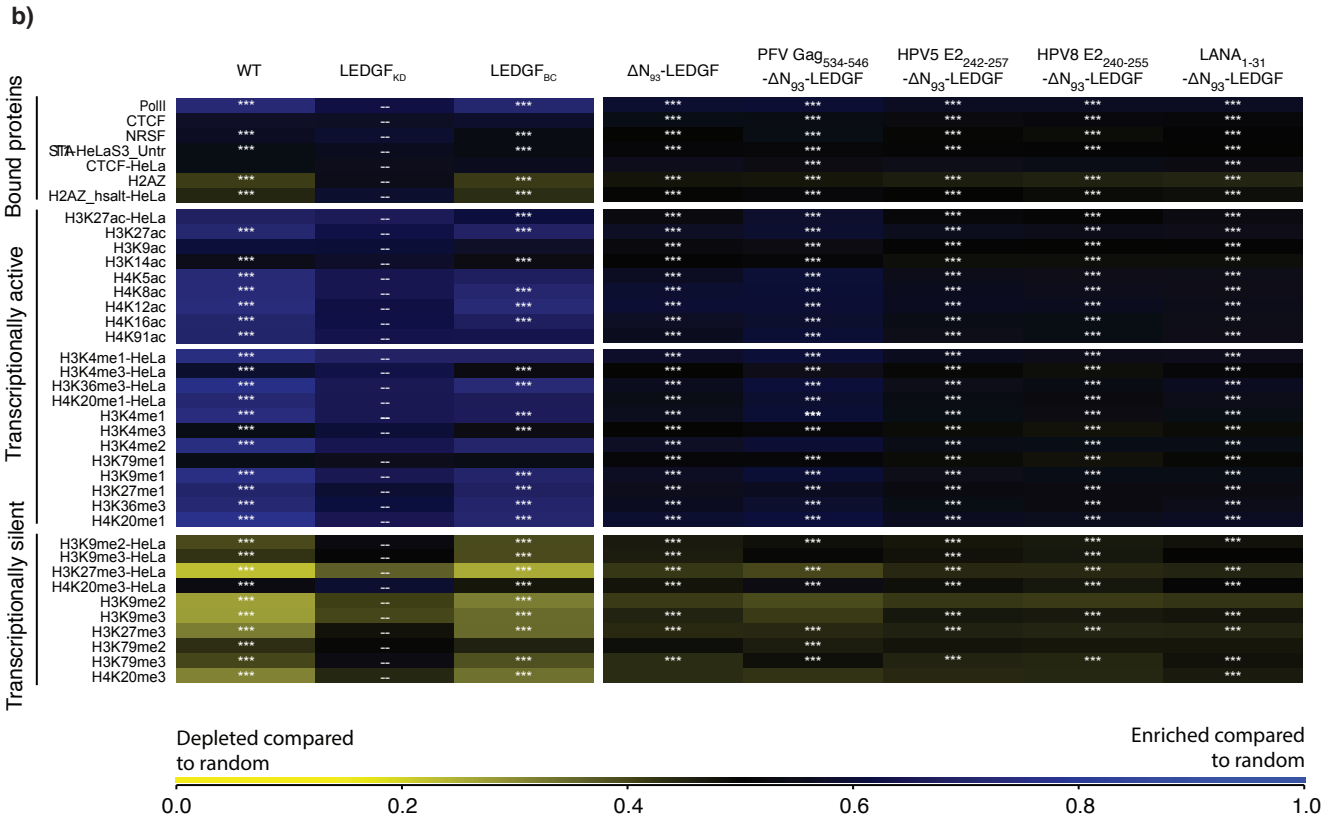


Figure 3.4: LEDGF-hybrids retarget lentiviral integration towards a more randomized pattern. (b) Epigenetic heat map comparing integration site data sets obtained from HeLaP4 LEDGF/p75 KD cells overexpressing different artificial LEDGF-hybrids to epigenetic features. Tile color depicting a positive or negative correlation to the respective epigenetic feature (10kb windows), relative to MRC, as indicated by the Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. Statistical significance (asterisks, ***p<0.001, ranked Wald tests) is shown relative to $LEDGF_{KD}$ population (dashed). Columns indicate different data sets while rows indicate different epigenetic features analyzed. Included features were limited to those identified in high-throughput studies HeLaP4 and primary $CD4^+$ T-cells. Detailed information on epigenetic marks and their roles can be found in [Barski et al., 2007; Taverna et al., 2007]. LANA, Latency Associated Nuclear Antigen; HPV, Human Papilloma Virus; PFV, Prototype Foamy Virus; LEDGF, Lens Epithelium-Derived Growth Factor

3.3.5 Artificial LEDGF/p75-peptide hybrids result in a safer integration profile

Together, the presented data above indicate that lentiviral vector integration preferences are defined by LEDGF/p75 as a cellular tether, and are mostly dictated by the N-terminal PWWP-domain. The mere deletion of this domain, or replacement with alternative chromatin-interacting modules redistributes vector integration sites in a more random fashion. The question remains whether redistribution of proviral integration sites obtained for our LEDGF-hybrids also translated in a safer therapy, with a lower chance on insertional mutagenesis. In an effort to get a better view on the safety profile, we calculated integration frequencies near a specific set of previously defined criteria [Papapetrou et al., 2011; Sadelain et al., 2011b], such as transcription start sites (<50kb), oncogenes (<300kb) or miRNA coding regions (<300kb), transcription units and ultraconserved elements to define potentially unsafe integra-

tion events. The large window sizes impose a very stringent selection for lentiviral integration events away from these features, which in turn can thus be considered as more safe [Papapetrou et al., 2011]. For each data set we evaluated the percentage of unsafe integrations (Table 3.3 and Supplementary Table S3.2 on page 56 and 67 respectively) and in addition determined the percentage of safe sites (events not captured in any of the other criteria; Table 3.3, % safe). When calculating the percentage in the parental cell line only 5.4% of all LV integration sites may be considered safe. LEDGF/p75-depletion results in a shift to 16.3% safe sites (p-value <0.005, Pearsons Chi-square compared to the $LEDGF_{WT}$ control condition), a phenotype that was fully reverted upon LEDGF/$p75_{BC}$ complementation (5.4%, no significant difference compare to $LEDGF_{WT}$). Ablation of the N-terminal PWWP-domain again boosted the percentage safe integration events to 19.7% (p-value <0.005, Pearsons Chi-square compared to the $LEDGF_{WT}$ control condition). Addition of heterologous peptide fragments KSHV $LANA_{1\text{-}31}$ and HPV8 $E2_{240\text{-}255}$ to the N-terminal end of $LEDGF_{93\text{-}530}$ slightly increased the % safe integrations relative to δN_{93}-LEDGF complementation up to 20.2% for HPV5 $E2_{242\text{-}257}$-δN_{93}-LEDGF, 21.2% for HPV8 $E2_{240\text{-}255}$-δN_{93}-LEDGF and 21.6% for KSHV $LANA_{1\text{-}31}$-δN_{93}-LEDGF when considering these criteria (p-value <0.005, Pearsons Chi-square compared to the δN_{93}-LEDGF control condition). Relative to the $LEDGF_{BC}$ condition our LEDGF-chimera increased the percentage of safe sites more than 3 fold (p-value <0.005, Pearsons Chi-square, Supplementary Table S3.2b). Of note, for the MRC conditions, we obtained a maximum of 30% integrations in safe harbors. Gene ontology (GO) analyses evaluating whether our different peptide fusions targeted genes belonging to different GO classes (subdivided based on (a) Biological processes and (b) Molecular function) did no show substantial shifts in integration frequencies for specific gene classes compared to those normally targeted by wild-type LEDGF (Supplementary Table S3.3 on page 68).

	Type	Controls	Total Sites	% TSS within 50kb		% Onco within 300kb		% miRNA within 300kb		% in transcript. Units		% in UCR		% Safe	
WT	insertion	TRUE	3520	36.0		43.2		28.0		79.4		7.0		5.4	
$LEDGF_{KD}$	insertion	FALSE	5560	33.8	**	33.8	***	21.4	***	57.0	***	6.2	*	16.3	***
$LEDGF_{BC}$	insertion	FALSE	1107	32.4		36.4		23.5		80.1		7.0		6.8	*
ΔN_{93}-LEDGF	insertion	FALSE	2990	29.7	***	31.1	***	20.4	***	56.5	***	6.5		19.1	***
PFV CBS $Gag_{534-546}$-ΔN_{93} LEDGF	insertion	FALSE	8267	33.5	**	35.4	***	22.4	***	57.0	***	5.9	**	17.4	***
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	insertion	FALSE	9572	30.3	***	31.4	***	19.7	***	54.2	***	6.1	*	20.5	***
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	insertion	FALSE	6616	30.0	***	31.8	***	20.4	***	54.0	***	6.0	*	21.2	***
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	insertion	FALSE	8565	30.6	***	32.9	***	21.6	***	54.2	***	5.5	***	21.6	***
WT	match	TRUE	10560	24.1		21.4		14.3		46.1		5.2		29.5	
$LEDGF_{KD}$	match	FALSE	16680	24.1		22.0		15.0		46.1		5.2		29.3	
$LEDGF_{BC}$	match	FALSE	3321	24.6		20.1		15.0		46.1		4.5		29.9	
ΔN_{93}-LEDGF	match	FALSE	8970	23.8		21.1		15.2		45.8		5.4		29.9	
PFV $Gag_{534-546}$-ΔN_{93}-LEDGF	match	FALSE	24801	24.3		21.1		14.4		45.7		5.1		29.9	
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	match	FALSE	28716	24.1		21.1		14.6		46.4		5.3		29.7	
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	match	FALSE	19848	24.2		21.2		14.7		46.2		4.8		29.5	
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	match	FALSE	25695	24.3		21.1		14.8		46.1		5.4		29.3	

Table 3.3: Integration frequency near safe harbor criteria. Table showing the percentage HIV-derived lentiviral vector integration frequencies near features (TSS, Oncogenes [Sadelain et al., 2011b], miRNA encoding regions, Transcription units and ultra conserved regions) that, when hit, are considered to be unsafe as defined in [Papapetrou et al., 2011] (Dataset details are described in the MM section). As such these features are used to define safe harbors as regions that fall outside these criteria. Percentages depict the fraction of integrations falling within the corresponding range relative to the criteria. The % of integrations negatively associated with these 5 features is used to calculate a safety profile. (*, p-value <0.5; **, p-value <0.05;***, p-value <0.005, Pearsons Chi-square compared to $LEDGF_{WT}$ control). TSS, Transcription Start Sites; UCR, Ultra Conserved Regions; PFV, Prototype Foamy Virus; HPV, Human Papilloma Virus; KSHV, Kaposi's Sarcoma Herpes Virus; LANA, Latency Associated Nuclear Antigen; match, matched random control.

3.4 Discussion

Integration of retroviral vectors into the host cell genome makes them invaluable tools for gene therapeutic applications where life-long correction is key. Different clinical trials showed effective gene transfer enabling long-term gene correction (For a review see [Naldini, 2015]). However, stable integration also implies the intrinsic risk of vector-induced genomic perturbation and the possibility of insertional mutagenesis leading to loss of function or transcriptional deregulation of neighbouring genes [Hacein-Bey-Abina et al., 2008; Howe et al., 2008; Cavazzana-Calvo et al., 2010]. These concerns became reality upon the emergence of leukemia in 25% of patients enrolled in a clinical trial for X-SCID and the development of acute lymphoblastic and myeloid leukemia in 70% of the patients in the Wiskott-Aldrich Syndrome (WAS) trial, both applying early generation gammaretroviral vectors [Boztug et al., 2010; Braun et al., 2014]. The yRV preference for integration into enhancer regions and concomitant activation of proto-oncogenes led to malignant transformation of cells and clonal expansion [Mitchell et al., 2004; Derse et al., 2007; Deichmann et al., 2007]. For therapeutic gene delivery attention has since turned to LVs due to their higher efficiency, safer integration profile and lower genotoxicity in preclinical models. In addition, improved vector design led to the development of self-inactivating vectors resulting in diminished enhancer activity of the U3 region and decreased genotoxicity of RV vectors [Zychlinski et al., 2008; Montini et al., 2009a; Modlich et al., 2009a; Newrzela et al., 2008]. Other attempts to limit the adverse events involved the use of insulator sequences as enhancer and silencer blockers [Emery, 2011], retargeting of viral integration [Ferris et al., 2010; Gijsbers et al., 2009; Silvers et al., 2010] and cell specific promoters to support cell specific expression [Antoniou et al., 2013]. However, RVs will always carry the intrinsic risk of open reading frame-disruption, as indicated by the report on SIN-LV affected splicing [Trono, 2012]. In addition, also LV integration may lead to clonal dominance as reported in the β-thalassemia trial [Cavazzana-Calvo et al., 2010]. Therefore it is important to gain additional mechanistic insights into the molecular mechanism of integration and integration site selection for LVs to be accepted for general therapeutic use. We and others substantially contributed to the elucidation of the role of LEDGF/p75 as a molecular tether of lentiviral vector integration. As a cellular cofactor of lentiviral integration, LEDGF/p75 orchestrates lentiviral integration preference by binding H3K36me$_3$ in the body of active transcription units via its N-terminal PWWP domain, but it is the vector-encoded integrase that catalyzes the integration reaction. Depletion of LEDGF/p75 by knockdown or knockout strategies shifts lentiviral vector integration out of active genes, yet integration is not completely random [Ciuffi et al., 2005; Schrijvers et al., 2012a], which at least in part can be explained by residual targeting via HRP-2 [Schrijvers et al., 2012a]. Here we set out to study whether different LEDGF-hybrids could be generated to distribute lentiviral integration sites more randomly. This line of vector development is based on the further increasing interest in new vector platforms displaying a close-to-random insertional profile potentially reducing the probability of proto-oncogene activation lowering the genotoxic potential [Staunstrup et al., 2009; Moldt et al., 2011; Suerth et al., 2012]. In an effort to achieve a more random integration site distribution, we deleted the specific chromatin-binding PWWP module of LEDGF/p75 (aa 1-93), or we replaced it with alternative pan-chromatin binding

modules. In case of LEDGF/p75, it is demonstrated that the PWWP domain recognizes H3K36me$_3$, a chromatin mark that is particularly enriched in the body of active transcription units [De Rijck et al., 2010; Pradeepa et al., 2012; Eidahl et al., 2013; Gijsbers et al., 2011a; van Nuland et al., 2013]. Complementation of LEDGF-depleted cells with a LEDGF/p75-protein that had its PWWP domain deleted (δN_{93}-LEDGF) or replaced with alternative chromatin binding modules showed unique subnuclear distributions for each of the constructs, indicating that these deletion of the PWWP domain, or the replacements with any of the other peptides, resulted in a specific redistribution within the nuclear compartment of the artificial LEDGF chimera (Figure 3.2 on page 47). The latter phenotype can be attributed to the AT-hook motifs and charged regions present in the N-terminal end of δN_{93}-LEDGF, together with the specific peptides that replaced the PWWP domain. After working up integration sites, analysis showed that lentiviral integration preferences for most of the constructs resulted in a more random distribution than under LEDGF depleted conditions (genomic and the epigenetic heat map representations; Figure 3.4a and b on page 53), except PFV Gag$_{534\text{-}546}$-δN_{93}-LEDGF. For example, in the latter cells LV integration was still enriched near epigenetic markers for transcriptionally active chromatin, albeit less outspoken than observed with LEDGF$_{WT}$ and LEDGF$_{BC}$ cells. Interestingly, peptide addition was not required to obtain a more random distribution. Lentiviral integrations in δN_{93}-LEDGF expressing cells were redistributed in a fairly random manner, with tile colors shifting to grey and black (for the genomic and the epigenetic heat map representations, respectively) indicating that integration frequencies for these features are not enriched nor depleted compared to the matched random integration site distribution. Comparison with LEDGF$_{KD}$ shows that integration is more randomly distributed than under LEDGF depletion (*** $p<0.001$, Wald statistics; Figures 3.4a, b and Supplementary Figure S3.2 and S3.3 on pages 64 and 65). Fusion of short pan-chromatin binding peptides to the truncated δN_{93}-LEDGF resulted in similar shifts towards a more randomized integration profile. The fact that all peptide fusions display a unique subnuclear location, suggest that their interaction with chromatin is different. Even though the overall integration frequencies are highly similar (considering the genomic and the epigenetic features analyzed), larger integration site datasets (>10e5 sites) would be required to allow more detailed analysis on the specific subsets. In an effort to estimate the effect of the more randomized distribution on safety, we calculated the frequency of integration relative to a set of safe harbor criteria for the individual integration site datasets [Papapetrou et al., 2011]. This analysis showed that the more random distributions resulted in a lower genotoxic profile with 18-22% of integrations meeting safe harbor criteria for our LEDGF-chimera compared to only 5.4% for cells carrying wild-type LEDGF/p75, all LEDGF-chimera resulted in a safer distributions over the genome. Fully targeted integration towards safe harbor regions like the AAVS1 or CCR5 locus would be the ultimate solution to circumvent insertional mutagenesis [Lombardo et al., 2011; Papapetrou et al., 2011; Sadelain et al., 2011b]. Several methods for site-directed gene correction have been developed using genetic scissors based on Zinc-finger nucleases, transcription activator like effector nucleases or more recently RNA-guided nucleases (CRISPR/Cas9) (for a review [Mussolino and Cathomen, 2013]). However, site directed integration would no doubt impair transduction efficiencies. Our approach improves the therapeutic potential of lentiviral vectors by decreasing the risk/benefit ratio, still sup-

porting high transduction efficiencies. The fact that integration can be directed to genomic regions that are not targeted under wild-type conditions nor LEDGF-depleted conditions, indicates that integration in these areas is disfavored due to the absence of a tether, rather than the presence of specific obstacles such as steric hindrance resulting from the condensed chromatin structure. As an alternative to the generation of stable cell lines as employed here, we demonstrated earlier that mRNA-electroporation ensures timely, high-level recombinant protein expression that is sufficient to retarget lentiviral vector integration [Vets et al., 2013]. When combined with IN mutant lentiviral vectors that selectively bind complementary LEDGF/p75 variants [Wang et al., 2014], this approach should be broadly applicable to introduce therapeutic or suicide genes for cell therapy, such as genetic modification of patient-specific iPS cells and improve safety of lentiviral vectors. With the occurrence of potential adverse effects being of multi-factorial nature [Kustikova et al., 2009] and assays lacking to predictively assess the cytotoxicity observed in vivo [Rothe et al., 2014b], a continuous effort aiming at abolishing the risk of insertional mutagenesis will be required for gene therapy to become a broadly accepted treatment alternative.

3.5 Experimental procedures

3.5.1 Generation of stable cell lines.

SIV-based vector transfer plasmids (pGAE) were a kind gift of D. Nègre (Laboratoire de Vectorologie Rétrovirale et Thérapie Génique, INSERM U412, IFR 74, Ecole Normale Supérieure de Lyon, Lyon, France). A lentiviral vector carrying CMV promoter driving a Zeocin resistance gene and a LEDGF specific miRNA-based shRNA was described earlier [Osório et al., 2014] and used to generate stable $LEDGF_{KD}$ cells. All LEDGF/p75 hybrid expression constructs were cloned into the pGAE backbone and cloning steps sequence verified.

Cloning of δN_{93}-LEDGF and δN_{93}-$LEDGF_{D366N}$ controls for the LEDGF δN_{93}-LEDGF fusions. pGAE_SFFV_ZnF4_δN_{93}_BC_I_BsdR_WPRE cl.9 and pGAE_SFFV_ZnF4_δN_{93}BC_D366N_I_BsdR_WPRE cl.3 were digested using *BglII* & *XhoI*. Ligation of the synthetic adaptor Ad_BglIIKO_AgeI_kozak generated pGAE_SFFV_ δN_{93}_BC_I_BsdR_WPRE cl. E and pGAE_SFFV_δN_{93}_BC_D366N_I_BsdR_WPRE cl. 5. We further refer to the controls as δN_{93}-LEDGF and δN_{93}-$LEDGF_{D366N}$ respectively.

Cloning of LEDGF δN_{93}-LEDGF and δN_{93}-$LEDGF_{D366N}$ hybrids. pGAE_SFFV_ZnF4_ δN_{93}_BC_I_BsdR_WPRE cl.9 and pGAE_SFFV_ZnF4_ δN_{93}BC_D366N_I_BsdR_WPRE cl.3 were digested using *BglII* & *XhoI*. Ligation of the synthetic adaptors (for adaptor sequences see Supplementary Table S3.1 on page 66) LANA31, PFVCBS13, HPV5E2_16 and HPV8E_216 generated:

pGAE_SFFV_LANA1-31_ δN_{93}_BC_I_BsdR_WPRE cl.A9
pGAE_SFFV_LANA1-31_ δN_{93}_BC_I_BsdR_WPRE cl.H
pGAE_SFFV_ PFVCBS13_ δN_{93}_BC_I_BsdR_WPRE cl.13
pGAE_SFFV_ PFVCBS13_ δN_{93}_BC_I_BsdR_WPRE cl.4
pGAE_SFFV_ HPV5E2_16 _ δN_{93}_BC_I_BsdR_WPRE cl.19

pGAE_SFFV_ HPV5E2_16 _ δN_{93}_BC_I_BsdR_WPRE cl.8
pGAE_SFFV_ HPV8E2_16 _ δN_{93}_BC_I_BsdR_WPRE cl.23
pGAE_SFFV_ HPV8E2_16 _ δN_{93}_BC_I_BsdR_WPRE cl.11

All cloning steps were confirmed by restriction digest and sequencing.

3.5.2 Cell culture.

All cells were grown in a humidified atmosphere containing 5 % CO_2 at 37 °C. HeLaP4 310 LEDGF/p75 depleted cells ([Gijsbers et al., 2009], further referred to as $LEDGF_{KD}$ cells) were grown in Dulbecco's Modified Eagle's Medium (DMEM; GIBCO-BRL, Merelbeke, Belgium) supplemented with 5% v/v heat inactivated Fetal Calf Serum (FCS; Sigma-Aldrich, Bornem, Belgium), 0.005% w/v gentamicin (GIBCO), 0.05% w/v geneticin (GIBCO) and 0.01% w/v zeocin (Life Technologies, Ghent, Belgium). These cells are monoclonal $LEDGF_{KD}$ cells, derived from HeLaP4 cells (gift from P. Charneau, Institut Pasteur, Paris, France). HelaP4 cells were grown on DMEM (GIBCO) supplemented with 5% v/v heat inactivated fetal calf serum (FCS; Sigma-Aldrich), 0.005% w/v gentamicin (GIBCO) & 0.05% w/v geneticin (GIBCO). HEK 293T cells (gift from O. Danos, Evry, France) were cultured in DMEM medium (GIBCO) with 8% v/v heat inactivated FCS (Sigma-Aldrich) and 0.005% w/v gentamicin (GIBCO). SupT1 cells were cultured in Roswell Park Memorial Institutes medium (RPMI, GIBCO-BRL, Merelbeke, Belgium) supplemented with 10% v/v heat inactivated fetal calf serum FCS (Sigma-Aldrich, Bornem,Belgium) and 0.005% w/v gentamicin (GIBCO). Nalm pre-B cells were cultured in RPMI (GIBCO) with 10% v/v heat inactivated FCS (Sigma-Aldrich) and 0.005% w/v gentamicin (GIBCO).

3.5.3 Retroviral vector production (SIV-based) and transduction.

Lentiviral vector production was performed as described earlier [Ibrahimi et al., 2009]. Briefly, for the generation of vesicular stomatitis virus glycoprotein (VSV-G) pseudo-typed SIV-based lentiviral vectors, HEK 293T cells were transfected with the packaging plasmid specific for SIV (pAd_SIV3+; gift from D. Nègre, Lyon, France), the envelope plasmid encoding VSV-G (pLP-VSVG #646 B, from Invitrogen) and respective transfer plasmids, using polyethylenimine (PEI; Polysciences, Amsterdam, The Netherlands). After collecting the supernatant, the medium was filtered using a 0.45 µm filter (Corning Inc., Seneffe, Belgium) and concentrated using a Vivaspin 15 50,000 MW column (Vivascience, Bornem, Belgium). The vector containing concentrate was then aliquoted per 50 µl and stored at -80 °C. Stable cell lines expressing a LEDGF-hybrid were generated by transduction of polyclonal LEDGF/p75 KD cells with SIV-based vectors and subsequent selection with 0,0003 % w/v blasticidin (Invitrogen). For lentiviral transduction experiments (LV eGFP T2A fLuc) cells were transduced over night (ON). 72 hours post-transduction cells were harvested when 90% confluent and used for eGFP FACS-analysis or measuring luciferase activity. The remainder of the transduced cells was further cultivated for at

least 10 days to eliminate non-integrated DNA and submitted for integration site sequencing.

3.5.4 Immunocytochemistry and Laser scanning microscopy.

Cells were transfected using Lipofectamine 2000 (Life Technologies, Merelbeke, Belgium) as described earlier [De Rijck et al., 2006]. LEDGF-hybrids were detected with the primary polyclonal rabbit anti $LEDGF_{480\text{-}530}$ antibody (A300-848a; 1/500; Bethyl Laboratories-Imtec Diagnostics N.V., Antwerpen, Belgium) and secondary polyclonal goat anti-rabbit antibody (1/500 in PBS, goat- αRb488; Bethyl Laboratories-Imtec Diagnostics N.V., Antwerpen, Belgium). Confocal images were acquired using an LSM 510 META imaging unit (Carl Zeiss, Zaventem, Belgium). Alexa-488 was excited at 488 nm (AI laser), mRFP at 543 (HeNe laser) and DAPI at 790 nm (Spectra-physics Mai Tai laser; Spectra Physics, Mountain View CA). After the main beam splitter (HFT KP 700/543 for mRFP, HFT UV/488/543/633 for eGFP, and HFT KP650 for DAPI) a secondary dichroic beam splitter was used to divide the fluorescence signal (NFT 490 for eGFP, NFT 545 for mRFP). Distinct signals were directed to different detectors and data analysis was performed with the LSM image browser. Overlay images were obtained using ImageJ freeware.

3.5.5 Western Blot.

Protein concentration of 1 % SDS (AppliChem, Leuven, Belgium) protein extracts sheared with a 27 G needle (Terumo, Leuven, Belgium) was determined using a bicinchoninic acid (BCA) protein assay (Pierce, Aalst, Belgium). Proteins were separated on a 12.5% w/v SDS-polyacrylamide gel and transferred to a polyvinylidene difluoride membrane (PVDF; BioRad) using an XCell SureLock electrophoresis system (Invitrogen). LEDGF-hybrids were detected using 1/2.000 polyclonal rabbit anti-$LEDGF_{480\text{-}530}$ antibody (A300-848a; Bethyl Laboratories-Imtec Diagnostics N.V., Antwerpen, Belgium) and 1/5 000 secondary antibody (polyclonal goat anti-rabbit antibody coupled with horse radish peroxidase (HRP); Dako). Chemiluminescence was measured using a ECL plus western blotting detection kit (Amersham Biosciences, Roosendaal, The Netherlands). Equal loading was verified with a primary monoclonal antibody directed to α-tubulin (mouse, 1/10 000, 1 h at room temperature; T5168, Sigma-Aldrich) and secondary antibody in blocking buffer (1/10 000, polyclonal goat-anti mouse labelled with HRP; Dako). Visualization was done by chemiluminescence (Pierce ECL Western Blotting Substrate, Thermo scientific).

3.5.6 Luciferase assay.

Cells were transduced with LV eGFP T2A fLuc and lysed with 70 µl of lysis buffer (50 mmol/l Tris pH 7.5, 200 mmol/l NaCl, 0.2% NP40, 10% glycerol). FLuc activity was determined using the ONE-glo luciferase assay system according to the manufacturers protocol (Promega, Leiden, The Netherlands) and normalized to the total protein concentration in order to correct for differences in metabolic state.

The total protein concentration was measured in parallel using a bicinchoninic acid (BCA) protein assay (Pierce, Aalst, Belgium).

3.5.7 Flow cytometric analysis.

Cells were transduced with LV eGFP T2A fLuc and harvested when 95% confluent. eGFP/YFP fluorescence was monitored by Flow cytometric analysis (FACS, Fluorescence activated cell sorting) using a FACSCalibur flow cytometer (BD Biosciences, Erembodegem, Belgium). Data analysis was performed with the CellQuest Pro software (BD Biosciences, Erembodegem-Aalst, Belgium). The percentage of eGFP positive cells (% of gated cells) multiplied by the mean fluorescence intensity (MFI) is further referred to as overall transduction efficiency.

3.5.8 Integration site amplification and sequencing.

Transduced cells were further cultivated for at least 10 days to eliminate non-integrated DNA. Cells were harvested when ca. 90% confluent. Genomic DNA was extracted using the GenElute Mammalian Genomic DNA miniprep kit (Sigma-Aldrich). Integration sites were amplified by linker-mediated PCR as described previously [Marshall et al., 2007] (see Supplementary Figure A.1 on page 163). Genomic DNA was digested using *MseI* and linkers were ligated (Supplementary Table S3.1 on page 66). Proviral-host junctions were amplified by nested PCR using Barcoded primers, generating 454 libraries. This enabled pooling of PCR products into one sequencing reaction. Products were gel-purified and sequenced using 454/Roche pyrosequencing (Titanium technology, Roche) on the 454 GS-FLX-instrument at the University of Pennsylvania. Reads were filtered based on perfect match to the LTR linker, Barcode and flanking LTR. All sites were mapped to the human genome requiring a perfect match within 3bp of the LTR end. Three Random Control sites were computationally generated and matched to experimental sites with respect to the distance to the nearest *MseI* Cleavage site (Matched Random Control, MRC). A more detailed explanation can be found in the supplementary guidelines of [Ocwieja et al., 2011]. Normalization of experimental HIV-derived lentiviral vector sites to those of the MRC sites functions as a control for recovery bias due to cleavage by restriction enzymes. Analysis was performed as described previously and genomic heat maps generated using the INSIPID software (Bushman Lab, University of Pennsylvania) [Marshall et al., 2007]. A detailed guide to interpret the data presented can be found in [Ocwieja et al., 2011]. The computation of DNase I site density was based on a table of DNase I sites obtained from [Lewinski et al., 2006]. Datasets used in the safe harbor analysis were retrieved from ENSEMBLE and/or UCSC (TxDB knownGenes, miRNA biotype, UCR; hg19) using BioMART [Kinsella et al., 2011]. The Allonco-list was used for oncogenes as published in [Sadelain et al., 2011b]. GO Slim classification analysis was done using the Web-based Gene set analysis toolkit (WebGeSTALT).

3.6 Author contributions

R.G. initiated design and concept of the study. L.V. conducted cell culture experiments. L.V., J.D. and R.G. analyzed the data. L.V, J.D, and R.G. performed bioinformatic analyses. L.V. and R.G. designed and cloned the different vectors. L.V., J.D., Z.D. and R.G. prepared the manuscript. Z.D. and R.G. were responsible for the coordination of the study. All authors read, corrected, and approved the final manuscript.

3.7 Acknowledgements

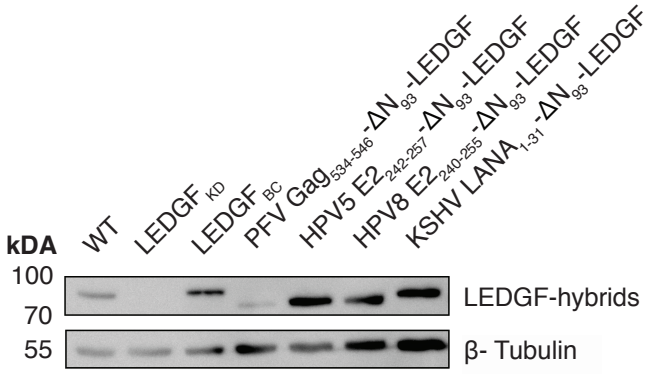
We are grateful to Jan De Rijck and Dominique Van Looveren for critical reading and to Paulien Van de Velde for her technical assistance. Viral vector production was performed at the Leuven Viral Vector Core. L.V. is a doctoral fellow supported by the Flemish Fund for Scientific Research (FWO; Fonds voor Wetenschappelijk Onderzoek). Research at KU Leuven received financial support from the FWO, the KU Leuven Research Council (OT), the KU Leuven IDO program and the Belgian IAP Belvir (IDO/12/008, ZKB9996 SB/0881057 and FWO ZKC0523, ZKC3378).

3.8 Disclosure statement

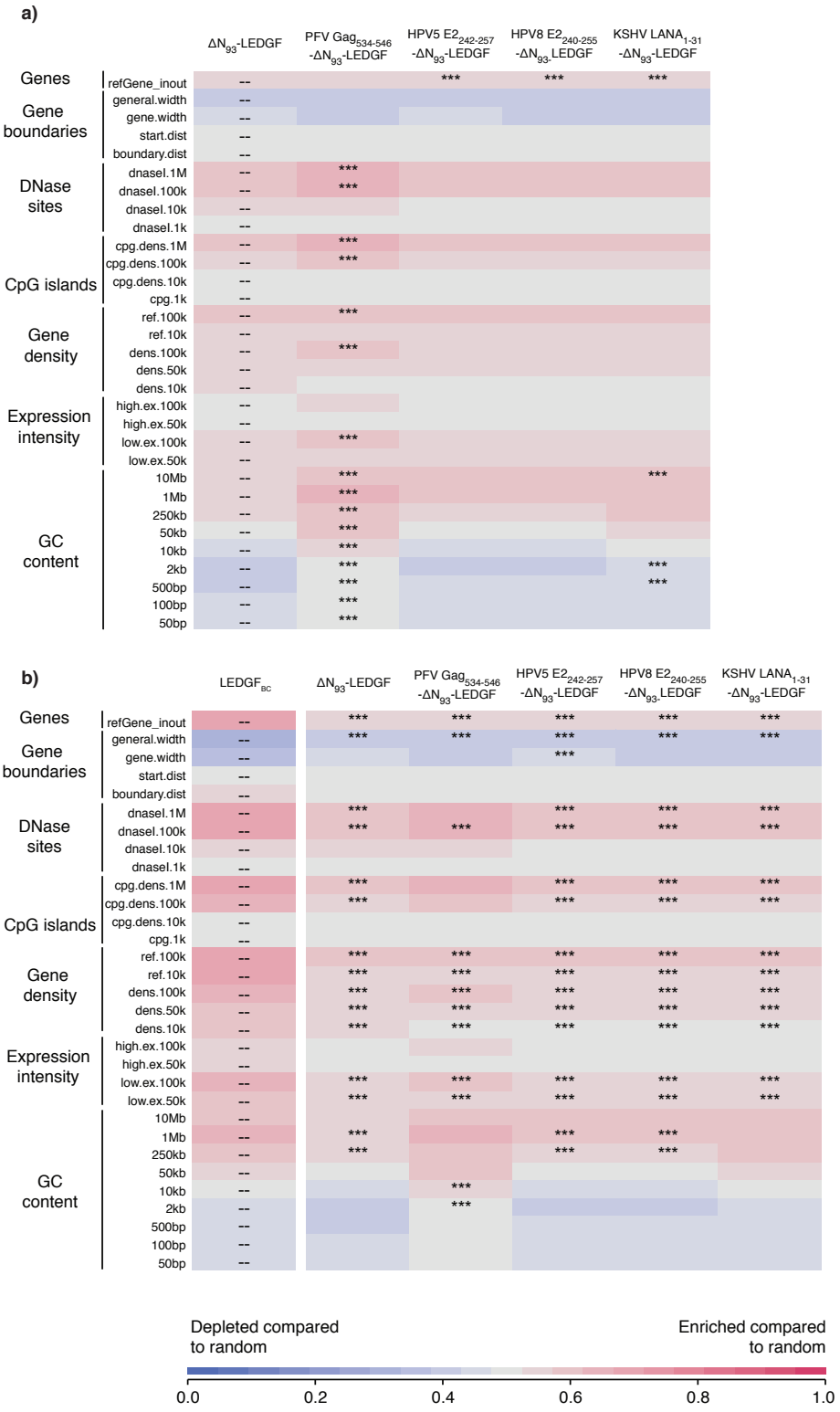
There are no conflicts of considerable interest.

3.9 Supporting information

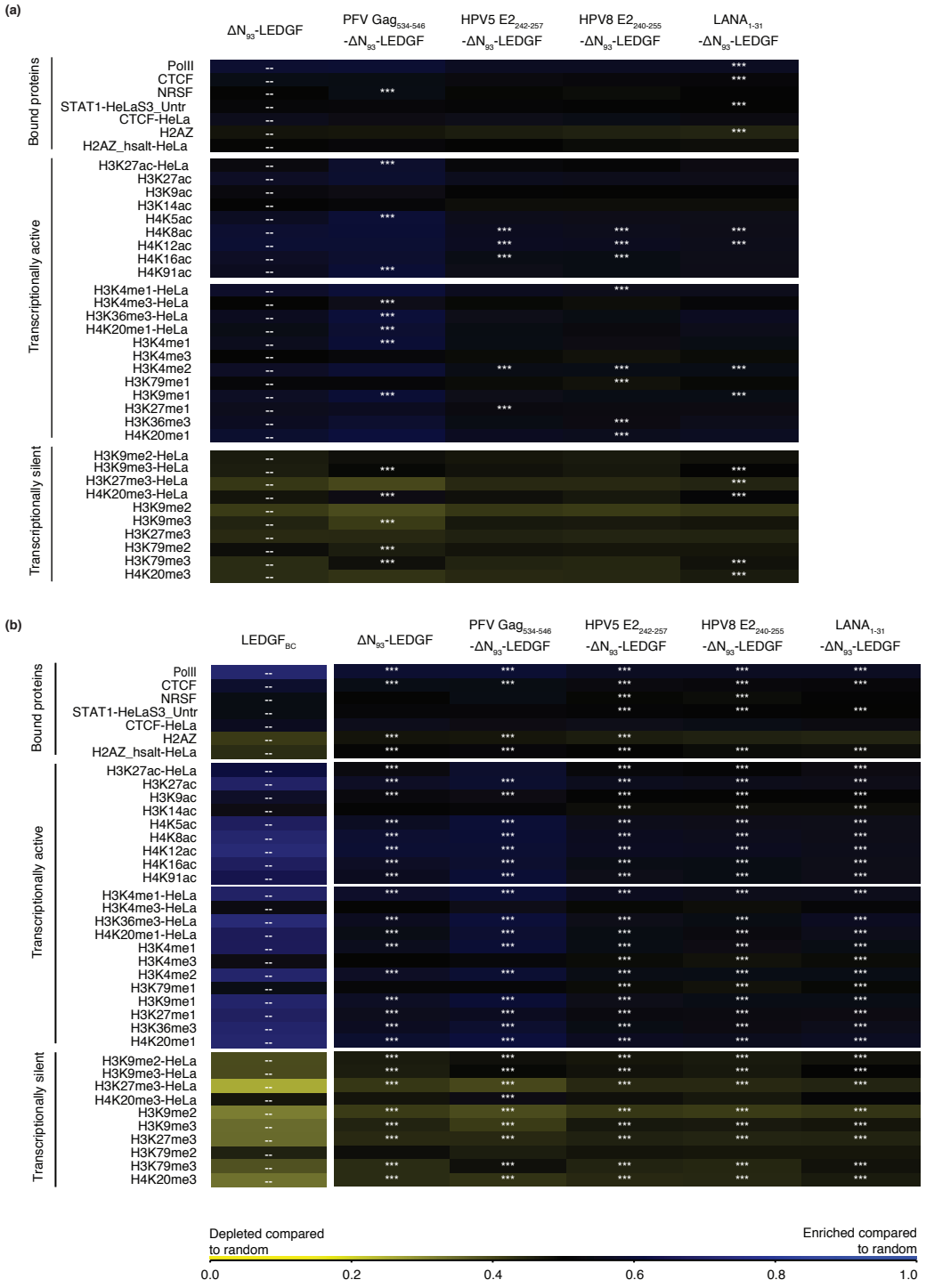
3.9.1 Supplementary figures



Supplementary Figure S3.1: Western analysis of LEDGF-fusions. LEDGF depleted cell lines were complemented with the respective LEDGF-hybrids. Total cell lysates were prepared and separated on a 12,5 % SDS gel. An antibody recognizing $LEDGF_{325\text{-}530}$ was used for detection. β-tubuline detection was used as an equal loading control. WT, Wild type; KD, KnockDown; PFV, Prototype Foamy Virus; LANA, Latency Associated Nuclear Antigen; HPV, Human Papilloma Virus; LEDGF, Lens Epithelium-Derived Growth Factor.



Supplementary Figure S3.2: LEDGF-hybrids retarget lentiviral integration towards a more randomized pattern. Genomic heat maps comparing integration site data sets obtained from HeLaP4 LEDGF/p75 KD cells overexpressing different artificial LEDGF-hybrids to genomic features. Tile color depicting the nature of the correlation for an integration dataset with the respective genomic feature (left) relative to matched random controls, as indicated by the colored Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. Statistical significance (asterisks, ***p<0.001, ranked Wald tests) is shown relative to (a) δN_{93}-LEDGF or (b) $LEDGF_{BC}$, respectively (double dash). Columns show different data sets while rows indicate different genomic features analyzed (described in [Ocwieja et al., 2011]). LANA, Latency associated nuclear antigen; HPV, Human Papilloma Virus; PFV, Prototype Foamy Virus; LEDGF, Lens Epithelium-Derived Growth Factor.



Supplementary Figure S3.3: LEDGF-hybrids retarget lentiviral integration towards a more randomized pattern. Epigenetic heat maps comparing integration site data sets obtained from HeLaP4 LEDGF/p75 depleted cells overexpressing different artificial LEDGF-hybrids to epigenetic features, generated using the INSIPID software (Bushman Lab, University of Pennsylvania). Tile color depicting a positive or negative correlation to the respective epigenetic feature (10 kb windows), relative to matched random controls, as indicated by the Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. Statistical significance (asterisks, ***p<0.001; ranked Wald tests) is shown relative to (a) δN_{93}-LEDGF or (b) $LEDGF_{BC}$, respectively (double dash). Significance is reached when p<0.001, compared to MRC. Columns indicate different data sets while rows indicate different epigenetic features analyzed. Included features were limited to those identified in high-throughput studies performed in HeLa and primary $CD4^+$ T-cells. Detailed information on epigenetic marks and their roles can be found in [Barski et al., 2007; Taverna et al., 2007]. LANA, Latency Associated Nuclear Antigen; HPV, Human Papilloma Virus; PFV, Prototype Foamy Virus; LEDGF, Lens Epithelium-Derived Growth Factor; MRC, Matched Random Control.

Name synthetic oligo	Sequence	Description
Ad_BglIIKO_AgeI_kozak_s	GATCACCGGTACCATGC	/
Ad_BglIIKO_AgeI_kozak_as	TGGCCATGGTACGAGCT	/
LANA31_s	GATCATGGCCCCTCCTGGCATGCGCCTGCGCAGCGGCCGCAGCACCGGCGCCCCTCTGACCCGCGGCAGCTGCCGCAAGCGCAACCGCAGCCCCGAGC	/
LANA31_as	TCGAGCTCGGGGCTGCGGTTGCGCTTGCGGCAGCTGCCGCGGGTCAGAGGGGCGCCGGTGCTGCGGCCGCTGCGCAGGCGCATGCCAGGAGGGGCCAT	/
PFVCBS13_s	GATCATGCAGGGCGGCTACAACCTGCGCCCCCGCACCTACCAGCCCC	/
PFVCBS13_as	TCGAGGGGCTGGTAGGTGCGGGGGCGCAGGTTGTAGCCGCCCTGCAT	/
HPV5E2_16_s	GATCATGCAGACCGAGACCCGCGGCCGCCGCTACGGCCGCCGCCCCAGCAGCAAGC	/
HPV5E2_16as	TCGAGCTTGCTGCTGGGGCGGCGGCCGTAGCGGCGGCCGCGGGTCTCGGTCTGCAT	/
HPV8E2_16_s	GATCATGCAGACCGAGACCAAGGGCCGCCGCTACGGCCGCCGCCCCAGCAGCCGCC	/
HPV8E2_16as	TCGAGGCGGCTGCTGGGGCGGCGGCCGTAGCGGCGGCCCTTGGTCTCGGTCTGCAT	/
Mse linker+	GTAATACGACTCACTATAGGGCTCCGCTTAAGGGAC	linker
Mse linker-	[Phosp]TAGTCCCTTAAGCGGAG-[AmC7-Q]	linker
MseL1	GTAATACGACTCACTATAGGGC	linker primer, PCR1
GW-3	gcctccctcgcgccatcagAGGGCTCCGCTTAAGGGAC	linker primer, PCR2, MseL2 fused to 454 primer A
MKL-3	CTTAAGCCTCAATAAAGCTTGCCTTGAG	HIV primer, PCR1
B-HIV3p-743	gccttgccagcccgctcagTCGTGATGagacccttttagtcagtgtggaaaatc	UPPERCASE IS BARCODE

Supplementary Table S3.1: Primer oligos used in this study. Table depicting the different oligos used in this study together with their nucleotide sequence.

(a)

	Type	Controls	Total Sites	% TSS within 50kb	% Onco within 300kb	% miRNA within 300kb	% in transcript. Units	% in UCR	% Safe
WT	insertion	FALSE	3520	36.0 ***	43.2 ***	28.0 ***	79.4 ***	7.0	5.4 ***
$LEDGF_{KD}$	insertion	FALSE	5560	33.8 ***	33.8 **	21.4 *	57.0	6.2	16.3 ***
$LEDGF_{BC}$	insertion	FALSE	1107	32.4 ***	36.4 ***	23.5 ***	80.1 ***	7.0 *	6.8 ***
ΔN_{93}-LEDGF	insertion	TRUE	2990	29.7 -	31.1 -	20.4 -	56.5 -	6.5 -	19.1 -
PFV CBS $Gag_{534-546}$-ΔN_{93} LEDGF	insertion	FALSE	8267	33.5 ***	35.4 ***	22.4 **	57.0	5.9 *	17.4 **
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	insertion	FALSE	9572	30.3	31.4	19.7 *	54.2 **	6.1 *	20.5 *
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	insertion	FALSE	6616	30.0	31.8	20.4	54.0 **	6.0 *	21.2 **
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	insertion	FALSE	8565	30.6 *	32.9 *	21.6 *	54.2 **	5.5 **	21.6 **
WT	match	FALSE	10560	24.1	21.4	14.3	46.1	5.2	29.5
$LEDGF_{KD}$	match	FALSE	16680	24.1	22.0	15.0	46.1	5.2	29.3
$LEDGF_{BC}$	match	FALSE	3321	24.6	20.1	15.0	46.1	4.5	29.9
ΔN_{93}-LEDGF	match	TRUE	8970	23.8	21.1	15.2	45.8	5.4	29.9
PFV $Gag_{534-546}$-ΔN_{93}-LEDGF	match	FALSE	24801	24.3	21.1	14.4	45.7	5.1	29.9
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	match	FALSE	28716	24.1	21.1	14.6	46.4	5.3	29.7
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	match	FALSE	19848	24.2	21.2	14.7	46.2	4.8	29.5
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	match	FALSE	25695	24.3	21.1	14.8	46.1	5.4	29.3

(b)

	Type	Controls	Total Sites	% TSS within 50kb	% Onco within 300kb	% miRNA within 300kb	% in transcript. Units	% in UCR	% Safe
WT	insertion	FALSE	3520	36.0	43.2	28.0	79.4	7.0	5.4 *
$LEDGF_{KD}$	insertion	FALSE	5560	33.8 **	33.8 ***	21.4 ***	57.0 ***	6.2 *	16.3 ***
$LEDGF_{BC}$	insertion	TRUE	1107	32.4 -	36.4 -	23.5 -	80.1 -	7.0 -	6.8 -
ΔN_{93}-LEDGF	insertion	FALSE	2990	29.7 ***	31.1 ***	20.4 ***	56.5 ***	6.5 *	19.1 ***
PFV CBS $Gag_{534-546}$-ΔN_{93} LEDGF	insertion	FALSE	8267	33.5 **	35.4 ***	22.4 ***	57.0 ***	5.9 **	17.4 ***
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	insertion	FALSE	9572	30.3 ***	31.4 ***	19.7 ***	54.2 ***	6.1 **	20.5 ***
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	insertion	FALSE	6616	30.0 ***	31.8 ***	20.4 ***	54.0 ***	6.0 **	21.2 ***
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	insertion	FALSE	8565	30.6 ***	32.9 ***	21.6 ***	54.2 ***	5.5 ***	21.6 ***
WT	match	FALSE	10560	24.1	21.4	14.3	46.1	5.2	29.5
$LEDGF_{KD}$	match	FALSE	16680	24.1	22.0	15.0	46.1	5.2	29.3
$LEDGF_{BC}$	match	TRUE	3321	24.6	20.1	15.0	46.1	4.5	29.9
ΔN_{93}-LEDGF	match	FALSE	8970	23.8	21.1	15.2	45.8	5.4	29.9
PFV $Gag_{534-546}$-ΔN_{93}-LEDGF	match	FALSE	24801	24.3	21.1	14.4	45.7	5.1	29.9
HPV5 $E2_{242-257}$-ΔN_{93}-LEDGF	match	FALSE	28716	24.1	21.1	14.6	46.4	5.3	29.7
HPV8 $E2_{240-255}$-ΔN_{93}-LEDGF	match	FALSE	19848	24.2	21.2	14.7	46.2	4.8	29.5
KSHV $LANA_{1-31}$-ΔN_{93}-LEDGF	match	FALSE	25695	24.3	21.1	14.8	46.1	5.4	29.3

Supplementary Table S3.2: Integration frequency near safe harbor criteria. Table showing the percentage HIV-derived lentiviral vector integration frequencies near features (TSS, Oncogenes [Sadelain et al., 2011b], miRNA encoding regions, Transcription units and ultra conserved regions) that, when hit, are considered to be unsafe as defined in [Papapetrou et al., 2011] (Dataset details are described in the MM section). As such these features are used to define safe harbors as regions that fall outside these criteria. Percentages depict the fraction of integrations falling within the corresponding range relative to the criteria. The % integrations negatively associated with these 5 features is used to calculate a safety profile. (*, p-value <0.5; **, p-value <0.05;***, p-value <0.005, Pearsons Chi-square compared to (a) δN_{93}-LEDGF or (b) $LEDGF_{BC}$ control condition). TSS, Transcription Start Sites; UCR, Ultra Conserved regions; PFV, Prototype Foamy Virus; HPV, Human Papilloma Virus; KSHV, Kaposi's Sarcoma Herpes Virus; LANA, Latency Associated Nuclear Antigen; match, matched random control.

a) Biological processes

	WT	$LEDGF_{KD}$	$LEDGF_{BC}$	ΔN_{93}-LEDGF	PFV $Gag_{534-546}$ -ΔN_{93}-LEDGF	HPV5 $E2_{242-257}$ -ΔN_{93}-LEDGF	HPV8 $E2_{240-255}$ -ΔN_{93}-LEDGF	KSHV $LANA_{1-31}$ -ΔN_{93}-LEDGF
Metabolic process	16.57	15.40	16.91	15.17	15.51	15.27	14.54	15.39
Biological regulation	15.00	14.64	15.05	14.83	14.40	14.49	14.54	14.99
Response to stimulus	10.79	11.10	10.83	11.42	10.62	10.90	10.93	10.64
Multicellular organismal process	8.77	9.47	8.74	7.88	9.51	9.31	9.57	9.57
Cellular component organization	8.57	7.81	8.47	8.44	7.54	8.21	7.80	7.77
Cell communication	8.17	8.41	7.87	8.82	8.39	8.50	8.56	8.41
Developmental pocess	7.63	8.06	7.82	7.57	7.92	7.96	8.05	7.87
Localization	7.51	6.81	6.94	7.47	7.39	7.19	7.43	7.46
Death	2.84	2.91	2.97	3.09	2.85	2.77	2.81	2.65
Reproduction	2.66	2.38	2.65	2.40	2.46	2.45	2.67	2.40
Cell proliferation	2.61	2.75	2.55	2.81	2.41	2.25	2.33	2.27
Multi-organism process	1.97	1.83	1.97	1.74	1.99	1.74	1.82	1.70
Growth	1.31	1.23	1.26	1.49	1.43	1.34	1.26	1.01
Unclassified	5.60	7.20	5.96	6.88	7.58	7.65	7.71	7.89

b) Molecular function

	WT	$LEDGF_{KD}$	$LEDGF_{BC}$	ΔN_{93}-LEDGF	PFV $Gag_{534-546}$ -ΔN_{93}-LEDGF	HPV5 $E2_{242-257}$ -ΔN_{93}-LEDGF	HPV8 $E2_{240-255}$ -ΔN_{93}-LEDGF	KSHV $LANA_{1-31}$ -ΔN_{93}-LEDGF
Protein binding	14.25	13.47	14.60	13.44	13.12	13.00	12.75	13.33
ion binding	10.89	10.32	10.83	10.94	11.27	11.12	10.88	11.02
Nucleic acid binding	6.16	5.00	6.00	4.69	5.33	5.19	4.90	5.76
Nucleotide binding	5.66	4.64	5.39	5.17	5.02	5.03	4.72	5.00
Hydrolase activity	4.05	3.96	4.45	4.17	4.44	4.29	4.18	4.31
Transferase activity	3.82	3.22	3.80	3.33	3.50	3.61	3.56	3.19
Enzym regulator activity	2.09	2.13	2.00	2.60	1.99	1.90	1.99	1.88
Transporter activity	1.58	1.72	1.37	1.63	2.13	1.84	2.10	2.04
Molecular transducer activity	1.39	1.93	1.43	2.26	1.99	1.91	2.14	2.12
Lipid binding	1.27	1.74	1.35	1.70	1.65	1.63	1.60	1.61
Chromatin binding	0.93	0.78	0.89	0.90	0.68	0.77	0.74	0.66
Structural molecule activity	0.83	0.72	0.85	0.90	0.83	0.67	0.82	0.87
Molecular adaptor activity	0.29	0.37	0.32	0.31	0.25	0.29	0.32	0.40
Carbohydrate binding	0.23	0.20	0.27	0.31	0.30	0.37	0.40	0.36
Electron carrier activity	0.21	0.21	0.15	0.17	0.18	0.14	0.12	0.10
Translation regulator activity	0.10	0.08	0.05	0.07	0.01	0.02	0.03	0.03
Antioxidant activity	0.06	0.08	0.03	0.00	0.01	0.01	0.02	0.01
Unclassified	5.52	6.62	5.47	6.77	7.10	7.08	7.43	7.34

Supplementary Table S3.3: Gene ontology analysis. Table depicting the gene ontology results using the a) Biological process or b) Molecular Function classification system. Row names depict the different Biological processes or Molecular Functions respectively. Numbers depict the percentage of integrations in genes included in the respective classes for each individual dataset. Color scale is based on the highest and lowest value in the table. Low percentages are represented in red while high percentages are represented in blue. Table showing the percentage HIV-based lentiviral vector integration frequencies near features. GO Slim classification analysis was done using the Web-based Gene set analysis toolkit (WebGeSTALT).

4

LEDGIN-Mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV

This chapter has been previously published as an article manuscript in EBioMedicine (2016):

Vranckx, L.S., Demeulemeester, J., Saleh, S., Boll, A., Vansant, G., Schrijvers, R., Weydert, C., Battivelli, E., Verdin, E., Cereseto, A., Christ, F., Gijsbers, R., Debyser, Z. (2016). LEDGIN-Mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV. EBioMedicine, Volume 8 , 248 - 264.

A patent application entitled "Functional Cure of Retroviral Infection" was filed to the UK Patent Office with myself (Lenard Sebastiaan Vranckx), Rik Gijsbers, Frauke Christ and Zeger Debyser listed as the inventors.

4.1 Abstract

Persistence of latent, replication-competent Human Immunodeficiency Virus type 1 (HIV-1) provirus is the main impediment towards a cure for HIV/AIDS (Acquired Immune Deficiency Syndrome). Therefore, different therapeutic strategies to eliminate this latent reservoir are currently being explored. We here propose a novel strategy to reduce the functional HIV reservoir during primary HIV infection by means of drug-induced retargeting of HIV integration. A novel class of integration inhibitors, referred to as LEDGINs, inhibit the interaction between HIV integrase and the LEDGF/p75 host cofactor, the main determinant of lentiviral integration site selection. We show for the first time that LEDGF/p75 depletion hampers HIV-1 reactivation in cell culture. Next, we demonstrate that LEDGINs relocate and retarget HIV integration resulting in an HIV reservoir that is refractory to reactivation by different latency-reversing agents. Taken together, these results support the potential of drugs that modulate integration site targeting to reduce the likeliness of viral rebound.

4.2 Research in context (Layman's terms)

Different strategies to cure HIV infection are being explored. Although complete eradication of the HIV provirus is the ultimate goal, disease remission allowing treatment interruption without viral rebound would constitute a significant leap forward. HIV integration site selection is orchestrated by LEDGF/p75. The advent of LEDGINs, that block the interaction between integrase and LEDGF/p75, allowed us to examine the hypothesis that interference with HIV integration site selection would yield integration sites that are less optimal for productive infection. Here we provide evidence in cell culture that LEDGIN treatment during acute HIV infection yields an HIV reservoir refractory to reactivation.

4.3 Introduction

Combination AntiRetroviral Therapy (cART) has revolutionized the treatment of HIV/AIDS, turning a life-threatening disease into a chronic illness. Yet, current therapies fail to cure infection due to the existence of a reservoir of latently infected cells [Archin et al., 2014b; Bruner et al., 2015; Siliciano and Siliciano, 2015]. While the Human Immunodeficiency Virus type 1 (HIV-1) actively replicates in activated $CD4^+$ T lymphocytes, it is able to reside in a long-lived quiescent state, mainly in resting memory $CD4^+$ T cells [Chun et al., 1997a,b; Finzi et al., 1997]. This latently infected cell population is established early on during infection and consists of a small fraction of the resting $CD4^+$ T cells in patients (about 1 in 10e6 cells) [Chun et al., 1997a,b; Finzi et al., 1997]. The reservoir enables HIV persistence during cART and is responsible for the rebound of viraemia upon therapy cessation [Richman et al., 2009]. Lentiviruses, such as HIV-1, preferentially integrate into transcriptionally active units [Schroder et al., 2002]. The latter integration preference is retained in latently HIV-1 infected primary $CD4^+$ T cells from patients [Han et al., 2004; Liu et al., 2006; Shan et al., 2011] and determined by Lens Epithelium-Derived Growth Factor (LEDGF/p75), a host-cell cofactor binding

HIV-1 IN via its C-terminal protein binding domain (Integrase Binding Domain (IBD)) [Cherepanov et al., 2003, 2005] and reading chromatin through its PWWP domain [Pradeepa et al., 2012; Eidahl et al., 2013]. LEDGF/p75 depletion shifts lentiviral integration out of transcription units [Ciuffi et al., 2005], a phenotype even more pronounced in human LEDGF/p75 KnockOut (KO) cells [Schrijvers et al., 2012a; Shun et al., 2007; Fadel et al., 2014]. In the absence of LEDGF/p75, its paralogue HRP-2 can at least in part take over this targeting role for HIV integration [Schrijvers et al., 2012b,a; Wang et al., 2012]. Structure-based drug design targeting the well-defined interface [Cherepanov et al., 2005] between the IBD and the HIV-1 IN catalytic core resulted in the development of 2-(quinolin-3-yl)acetic acid derivatives that inhibit HIV-1 replication [Christ et al., 2010]. This novel class of antivirals is referred to as LEDGINs [Christ and Debyser, 2013; Christ et al., 2010, 2012; Debyser et al., 2015; Demeulemeester et al., 2014a]. Novel congeners with nanomolar activity act as allosteric inhibitors, preventing the binding of both LEDGF/p75 and HRP-2 and interfering with the catalytic activity of IN [Christ et al., 2012; Kessl et al., 2012; Tsiang et al., 2012]. Recently, LEDGINs were found to inhibit late stage HIV replication as well [Jurado et al., 2013; Desimmie et al., 2013; Balakrishnan et al., 2013; Le Rouzic et al., 2013]. The phenotype requires the binding of LEDGINs to the LEDGF/p75 binding pocket in IN [Desimmie et al., 2013; Le Rouzic et al., 2013] and is mediated by enhanced multimerisation of IN in the viral particles [Jurado et al., 2013; Desimmie et al., 2013; Balakrishnan et al., 2013; Borrenberghs et al., 2014]. Inconclusive results have been obtained as to whether LEDGINs affect the integration site distribution [Sharma et al., 2014; Gupta et al., 2014]. In any case LEDGIN treatment results in a steep dose-dependent inhibition of viral replication in cell culture, supporting their clinical development [Fader et al., 2014; Fenwick et al., 2014]. Here we investigated the early effect of LEDGINs and evaluated their effect on HIV integration site distribution. In addition, we monitored the effect of LEDGINs on the establishment of the latent reservoir and investigated whether retargeting of integration could lead to a silent HIV reservoir resistant to reactivation. In a stepwise approach we first demonstrate that LEDGF/p75 depletion results in reduced integration but a relatively expanded silent HIV reservoir. Next, we demonstrate that upon treatment with LEDGINs, blocking the LEDGF/p75-IN interaction, the residual proviral integration shifts away from transcription units. LEDGIN treatment also shifts the 3D localization of the integrated provirus towards the inner nucleus. LEDGIN-induced retargeting relatively expands the silent HIV reservoir in cell lines and primary $CD4^+$ cells. This silent reservoir is refractory to reactivation by latency reversing agents (LRAs). Pushing sufficient proviruses into latency is theoretically predicted to drive the basic reproduction number of HIV below 1, resulting in an unsustainable infection [Rouzine et al., 2015]. Hence, addition of LEDGINs to cART regimens during acute HIV infection may represent a new strategy to achieve a remission of HIV infection in patients.

4.4 Results

4.4.1 LEDGF/p75 depletion results in a quiescent reservoir.

We and others reported on retargeted proviral integration in LEDGF/p75-depleted cells [Ciuffi et al., 2005; Schrijvers et al., 2012a; Shun et al., 2007; Fadel et al., 2014]. To study the role of LEDGF/p75 in establishing the latent reservoir, we used a variant of the recently developed double reporter virus that simultaneously measures a constitutive and a LTR-driven reporter (see methods, Figure 4.1a on page 73, [Calvanese et al., 2013; Chavez et al., 2015]). This Orange-Green HIV-1 (OGH) single-round reporter virus carries an LTR-driven enhanced Green Fluorescent Protein (eGFP) together with a constitutively active EF1α promoter driving monomeric Kusabira-Orange2 (mKO2) expression (Figure 4.1a). The double-fluorescent reporter virus allows quantification of distinct populations in the infected cell pool via FACS analysis (Figure 4.1b). Through the constitutively active EF1α promoter all infected cells express the mKO2 reporter. Provirus with an active LTR also expresses the eGFP reporter and is referred to as double positive, active virus, whereas provirus with a quiescent LTR does not express eGFP and is called here upon the quiescent provirus. Wild type and LEDGF/p75-depleted SupT1 cells (Supplementary Figure S4.1a, bottom panel, on page 102) were infected with a dilution series of single-round OGH virus. The percentage of infected cells was evaluated at three days post infection, discriminating productively infected cells (active provirus; $eGFP^+$, $mKO2^+$) from the latently infected population (quiescent provirus; $eGFP^-$, $mKO2^+$). Single reporter constructs were used for validation (Figure 4.1b). As expected, LEDGF/p75-depletion reduced HIV-1 infection 2- to 3-fold as judged by the percentage mKO2 positive cells (Supplementary Figure S4.1a), in line with earlier data [Vandekerckhove et al., 2006; Gijsbers et al., 2009; Schrijvers et al., 2012a]. Evaluation of the percentage of $eGFP^-$, $mKO2^+$ cells relative to the total number of infected cells ($mKO2^+$) provides an estimate of the fraction of quiescently infected cells in the infected pool (% $eGFP^-$, $mKO2^+$ cells) / (% $mKO2^+$ cells) *100. At all virus dilutions tested, LEDGF/p75 depleted cells contained more quiescent proviruses than WT cells (Figure 4.1c). When this experiment was repeated in the $Nalm^{+/c}$ control and $Nalm^{-/-}$ LEDGF/$p75_{KO}$ cell lines [Schrijvers et al., 2012a], similar effects were observed (Supplementary Figure S4.1a, b). Since LEDGF/p75 depletion results in redistribution of HIV integration [Ciuffi et al., 2005; Schrijvers et al., 2012a; Shun et al., 2007; Fadel et al., 2014], these results suggest that the altered integration site distribution after LEDGF/p75 depletion increases the transcriptionally quiescent fraction.

4.4.2 LEDGF/p75 depletion decreases the reactivation potential of the quiescent reservoir.

To further characterize the latently infected cell pool that is generated in LEDGF/p75-depleted cells, we set out to reactivate the latent provirus using different Latency Reversing Agents (LRAs). Here, we used the LEDGF/$p75_{KO}$ cell line ($Nalm^{-/-}$) to study whether latent provirus in LEDGF/p75-depleted cells has an altered reactivation potential. We made use of a NL4.3-based single reporter virus containing

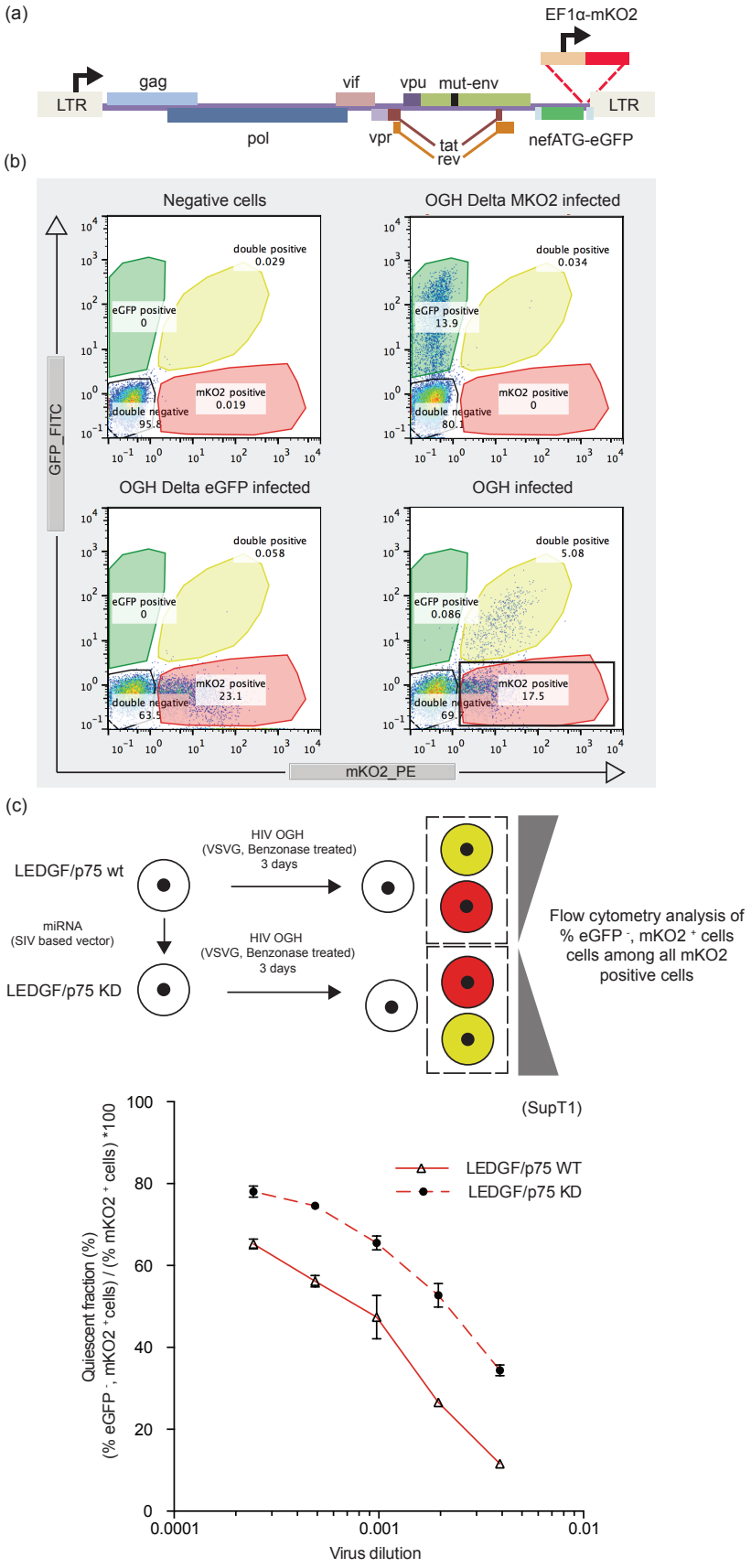


Figure 4.1: LEDGF/p75 depletion increases the silent reservoir.

Figure 4.1: LEDGF/p75 depletion increases the silent reservoir. (a) Schematic representation of the two-colored reporter virus carrying an eGFP driven by the viral LTR promoter in the nef position and an entire constitutive transcriptional unit (EF1α-mKO2) inserted downstream. (b) Dot plots representing FACS analysis of SupT1 cells infected with the single reporter viral controls (OGH-δmKO2, OGH-δeGFP) or the double reporter virus OGH. The different cell populations are highlighted in the representative color. (c) $LEDGF_{KD}$ affects the fraction of silently infected cells (% $eGFP^-$, $mKO2^+$ cells) / (% $mKO2^+$ cells) *100. Data represent averages of triplicates from a representative experiment and error bars indicate the standard deviation. All viruses are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein; mKO2, Mutant Kusubira Orange 2.

an LTR-driven truncated CD34 (tCD34) as a reporter protein (NL4-3.tCD34.R-.E-/VSV-G, Figure 4.2a on page 75). Control ($Nalm^{+/c}$) and $LEDGF/p75_{KO}$ ($Nalm^{-/-}$) cells were infected with a dilution series of HIV-tCD34. Virus dilutions were selected to result in equal integrated copy numbers, both the LEDGF/p75 control ($Nalm^{+/c}$) (1.45E-01 ± 1.40E-02 copies) and $LEDGF/p75_{KO}$ ($Nalm^{-/-}$) cells (1.45E-01 ± 2.41E-02 copies) to compare the fold reactivation between both. Cells were reactivated 11 days post infection with different LRAs (Figure 4.2b), and the fold reactivation (% tCD34) relative to DMSO between $LEDGF/p75_{WT}$ and $LEDGF/p75_{KO}$ conditions was measured 24 hrs later. The percentage living cells amounted at least 78% of the total cell population in all conditions and was independent of LEDGF/p75 depletion (data not shown). Modest tCD34 reactivation was observed after addition of Phorbol Myristate Acetate (PMA) or Prostratin. An increase in the percentage tCD34 positive cells of 1.5-2 fold was observed for the $Nalm^{+/c}$ control while $LEDGF/p75_{KO}$ ($Nalm^{-/-}$) conditions only experienced an increase of 1.3-1.5 fold when adding PMA (0.3-3 µM) or Prostratin (5 µM) (t-test with Sidak-Bonferroni correction; * $p<0.05$, control compared to $LEDGF/p75_{KO}$). A similar effect was observed when stimulating with 1 µM SuberoylAnilide Hydroxamic Acid (SAHA). Yet, addition of 3 µM of SAHA resulted in a 3.2-fold reactivation of tCD34 in the presence of LEDGF/p75 but only a 2.2-fold reactivation in its absence (t-test with Sidak-Bonferroni correction; ** $p<0.005$, WT compared to $LEDGF/p75_{KO}$) (Figure 4.2b). Together, these data indicate that integration in the absence of LEDGF/p75 results in a larger quiescent cell pool upon infection (Figure 4.1c) with a relatively larger proportion of cells refractory to reactivation (Figure 4.2b).

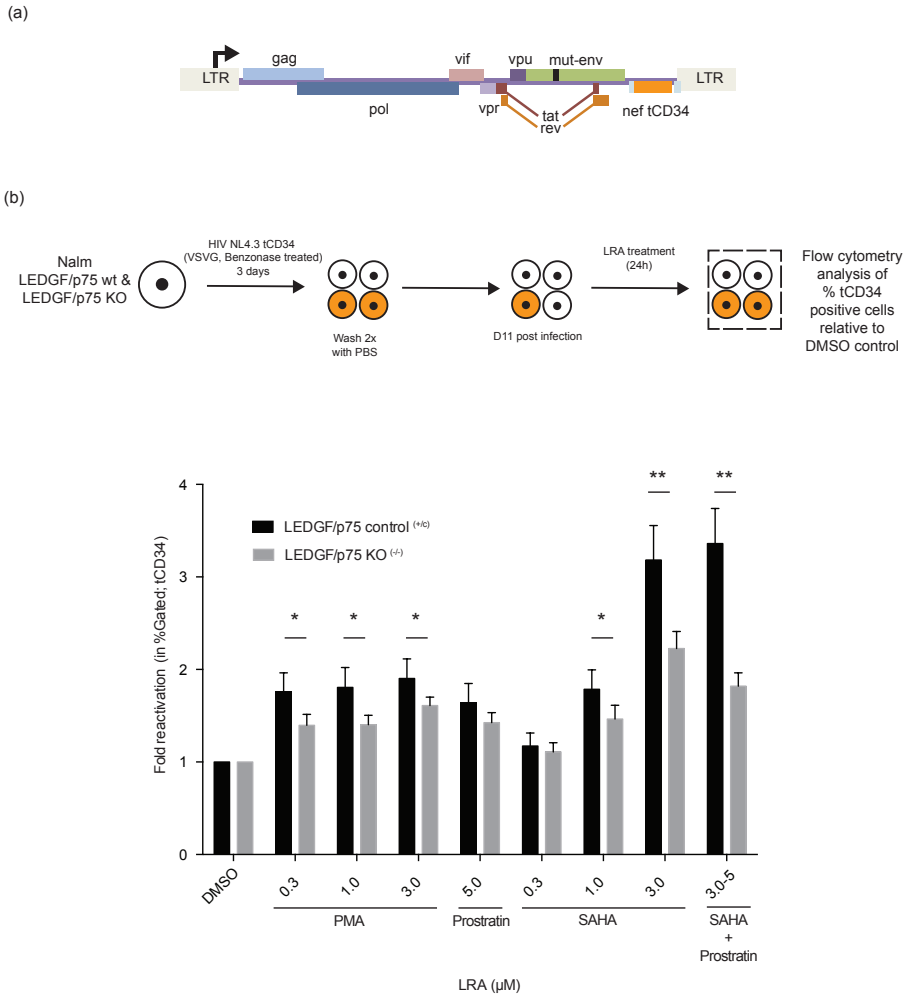


Figure 4.2: LEDGF/p75 depletion reduces HIV reactivation from latency. (a) Schematic representation of the single round HIV reporter virus encoding a tCD34 driven by the viral LTR promoter in the nef position. (b) Bar diagram depicting the fold reactivation (as fold increase in % tCD34 positive cells). Nalm control (+/c) and Nalm LEDGF/p75$_{KO}$ (-/-) cells were infected with a dilution series of single round reporter virus and the % tCD34 positive cells was monitored. 11 days post infection cells were reactivated using different LRAs at the concentrations indicated. Data represent averages of 9 replicates from 3 independent experiments and error bars indicate the standard error of the mean (SEM). A statistical analysis was performed using multiple t tests and corrected using Sidak-Bonferroni (* $p < 0.05$, ** $p < 0.005$ vs. LEDGF/p75$_{KO}$) (SAHA; SuberoylAnilide Hydroxamic Acid, PMA; Phorbol 12- Myristate 13-Acetate, Prostratin, DMSO; DiMethyl SulfOxide.). Normalization was based on equal integrated copy (IC) numbers. All viruses are VSV-G pseudotyped. tCD34; truncated Cluster of Differentiation 34.

4.4.3 LEDGIN treatment shifts HIV integration out of transcription units.

Next, we evaluated the effect of recently developed LEDGINs, small molecules that inhibit LEDGF/p75-IN interaction and HIV integration [Christ et al., 2010; Demeulemeester et al., 2014a], on the HIV reservoir in cell culture. SupT1 cells were transduced with a single-round HIV-based lentiviral vector expressing eGFP in the presence of a dilution series of LEDGIN CX014442 [Christ et al., 2012]. Flow cytometry and Q-PCR revealed a dose-dependent decrease in lentiviral transduction as represented by the % eGFP-positive cells and the number of integrated copies (Supplementary Figure S4.2a and

b on page 104). In a first step, we determined the distribution of HIV-based viral vector integration sites [Marshall et al., 2007; Gijsbers et al., 2009]. The number of integration sites is indicated for each data set (Table 4.1 and Figure 4.3 on page 78 and 79 respectively). We analysed lentiviral integration frequencies relative to a set of genomic features (Table 4.1). In line with previous results, HIV integration in WT SupT1 cells was enriched in the body of genes (69.54% in RefSeq genes (Table 4.1)) disfavoring transcription start sites (TSS) and promoter regions (1.78% with in 2kb of the 5' of a RefSeq gene and 2.02% within 2kb of a CpG island) [Schroder et al., 2002; Mitchell et al., 2004]. The integration sites shifted out of transcription units under LEDGIN-treatment (54.55% in RefSeq genes (50 µM); *** $p<0.0001$, Chi-square test compared to DMSO) while integration close to TSS (6.94% (50 µM); *** $p<0.0001$, Chi-square test compared to DMSO) and CpG islands (5.50% (50 µM); ** $p<0.01$, Chi-square test compared to DMSO) increased in a dose-dependent manner. This shift in integration site distribution surpasses the one observed in LEDGF/p75-depleted cells [Ciuffi et al., 2005; Shun et al., 2007; Schrijvers et al., 2012a; Fadel et al., 2014]. The fact that LEDGINs also inhibit the interaction between HRP-2 and HIV-1 IN [Schrijvers et al., 2012a], can explain this observation. Comparable data were observed for larger window sizes (2kb and 4kb are shown). Our results were corroborated in MT4 cells using multiple round (WT) HIV NL4-3 and using the less potent LEDGIN CX05045 (Supplementary Table S4.1 on page 113) [Christ et al., 2010]. A genomic heatmap comparing integration site data sets obtained in SupT1 LEDGF/$p75_{KD}$ cells with WT SupT1 cells in the presence of various concentrations of LEDGIN is shown in Figure 4.3a. Analysis of global integration preferences clearly indicates a shifts out of transcriptionally active regions upon LEDGF/$p75_{KD}$ (Figure 4.3a, compared to DMSO), in line with previously reported data [Shun et al., 2007; Marshall et al., 2007; Gijsbers et al., 2009]. A similar shift was also observed under LEDGIN treatment at concentrations above 6 µM. In a more elaborate analysis we analyzed integration site frequencies relative to epigenetic features described in T cells (Supplementary Figure S4.3b & c on page 106).

Under WT conditions, HIV integration preferentially occurred near epigenetic markers associated with transcriptionally active regions (H3K4 mono-, di- and tri methylation, H3K14 and H4 acetylation, as well as acetylation or mono-methylation of H3K9/K27/K79, H4K20 and H2BK5,...) [De Ravin et al., 2014], while integration in transcriptionally silent regions or heterochromatin is disfavored ($H3K27me_3$, $H3K9me_3$ or $H4K20me_3$ and $H3K79me_3$, respectively). The overall integration profile is closer to random upon addition of LEDGIN CX014442 (as shown by the decrease in color intensity towards black, cfr. Color key Supplementary Figure S4.3b and c). Supplementary Figure S4.3c displays a more condensed heat map where epigenetic marks are grouped according to the respective chromatin states they associate with.

Since LEDGINs may potentially affect the inherent integration mechanism, resulting in aberrant integration or an altered local integration site preference, we evaluated sequence conservation and relative base frequency in the 18 bp genomic DNA sequence surrounding the integration sites (corresponding to the intasome footprint) using sequence logos (Supplementary Figure S4.5 on page 110). The WT local palindromic sequence logo surrounding the integration site [Holman and Coffin, 2005; Demeulemeester et al., 2014b]) was maintained in all conditions. In addition we calculated the percentage unique inte-

gration sites with an imperfect LTR-chromosome junction relative to the total integration sites after LEDGIN treatment (Supplementary Figure S4.6 on page 111 and Supplementary Table S4.2 on page 113, Supplementary experimental procedures). Only a minor fraction (<2%) of the total integration sites was identified as containing imperfect LTR-chromosome junctions. We conclude that although LEDGINs retarget integration towards more random, with the residual integration events representing authentic integration without gross LTR deletions.

(SupT1)

DataSet	Type	controls	TotalSites	InRefGene	% TSS within 2kb	% TSS within 4kb	% CpG within 2kb	% CpG within 4kb	% DHS within 2kb	% DHS within 4kb
DMSO ctrl	insertion	TRUE	3312	69.54	1.78	5.62	2.02	6.16	23.04	39.86
.5uM	insertion	FALSE	2451	70.58	2.00	5.92	2.33	5.83	24.07	39.94
1.5uM	insertion	FALSE	2278	70.37	2.41	5.88	2.37	5.79	20.94	37.18
3uM	insertion	FALSE	2485	66.60	1.97	5.59	1.73	5.39	21.49	37.55
6uM	insertion	FALSE	3365	65.23 **	3.03	6.45	2.97	6.39	21.78	37.12
12uM	insertion	FALSE	884	61.88 ***	3.96 **	7.58	3.39	7.35	23.64	38.57
25uM	insertion	FALSE	604	60.93 ***	4.30 **	8.94	4.97 **	10.26 **	22.19	40.07
50uM	insertion	FALSE	418	54.55 ***	6.94 ***	12.68 ***	5.50 **	11.72 ***	26.08	40.67
DMSO ctrl	match	TRUE	9918	39.12	2.15	5.09	2.33	5.04	15.08	26.72
.5uM	match	FALSE	7341	39.71	1.98	4.41	2.38	4.62	15.79	27.23
1.5uM	match	FALSE	6828	39.75	2.72	5.02	2.46	5.1	15.35	27.61
3uM	match	FALSE	7434	39.8	2.37	4.94	2.22	4.75	15.04	25.99
6uM	match	FALSE	10080	40.92	2.43	4.57	2.63	4.65	15.52	26.99
12uM	match	FALSE	2649	38.01	2.45	4.72	2.27	4.53	14.99	26.54
25uM	match	FALSE	1806	38.65	2.6	4.93	2.05	4.37	14.89	27.3
50uM	match	FALSE	1251	39.73	1.92	4.08	2.4	4.4	16.07	27.34

Table 4.1: Integration frequency near mapped genomic features in the human genome. Table showing the percentage of HIV-based vector integration sites relative to features specific for integration such as integration into the body of genes (Refseq genes), TSS, CpG islands and DNase I-hypersensitive sites. 2kb and 4kb windows are shown (data are obtained from SupT1 cells). Colors depicts the nature of the association based on the highest and lowest value in the column. Asterisks depict a significant deviation from the DMSO treated control dataset (two-tailed Chi-square test; ***, p-values <0.001). TSS, Transcription Start Sites; DHS, DNase I-Hypersensitive Sites. Colors depicts the nature of the association based on the highest and lowest value in the column.

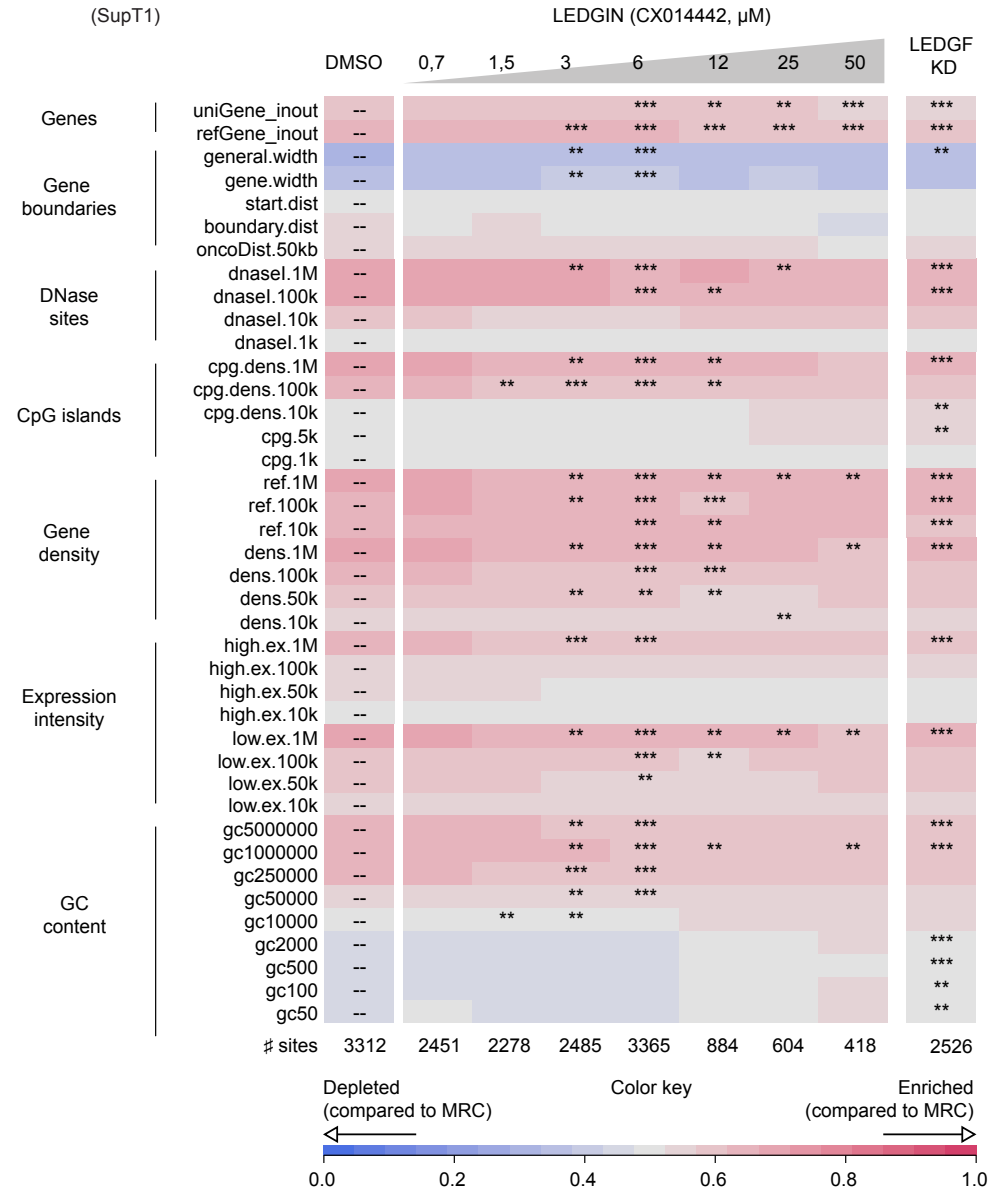


Figure 4.3: LEDGIN inhibition of the LEDGF/p75-IN interaction retargets lentiviral integration.

4.4.4 Abrogation of LEDGF/p75-IN interaction shifts 3D localization of the integrated provirus towards the inner nuclear compartment.

HIV-1 PICs and integrated HIV provirus preferentially localize in the nuclear periphery [Albanese et al., 2008; Di Primio et al., 2013; Francis et al., 2014]. Recent reports associate preferential integration with nuclear import and distance to nuclear pore complexes [Marini et al., 2015; Lelek et al., 2015]. Here we analyzed the 3D distribution of HIV-1 integrated provirus upon interruption of the LEDGF/p75-IN interaction. We first compared the distribution of HIV proviruses 48 h post infection in the Single Cell Imaging of Proviral HIV-1 assay (SCIP-assay) [Di Primio et al., 2013] between LEDGF/p75$_{WT}$ and LEDGF/p75 depleted U2OS cells (Figure 4.4a on page 81). The distribution of integrated HIV provirus

Figure 4.3: LEDGIN inhibition of the LEDGF/p75-IN interaction retargets lentiviral integration. Integration site data sets obtained from SupT1 cells infected with LVP2A and treated with different concentrations of CX014442 (IC_{50} CX014442 'early effect': 3.83 µM) were compared to different genomic features. A heat map was generated using the INSIPID software (Bushman Lab, University of Pennsylvania). Tile color is depicting the nature of the correlation for an integration data set with the respective genomic feature (rows, left) relative to matched random controls, as indicated by the colored Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. Columns indicate different data sets. Statistical significance (asterisks, ranked Wald tests) is shown relative to the DMSO population (dashes). Significance is reached when $p<0.001$, compared to the DMSO (**$p<0.01$; *** $p<0.001$). CX014442 treatment during infection shows a dose dependent shift out of transcriptionally active regions. Lower significance is observed at the highest LEDGIN concentrations due to a lower copy number of integration sites. (a more detailed guide to the data presented can be found in [Ocwieja et al., 2011]).

was analyzed by detection of γH2AX foci after I-Sce1 digestion (Figure 4.4a). Two-fold less integrants were detected after LEDGF/p75 depletion (Figure 4.4a). Whereas in WT cells the viral genomes localized near the nuclear rim (Figure 4.4b), the location of integrated provirus shifted towards the inner nuclear compartment after LEDGF/p75 depletion as indicated by the cumulative frequency plotted relative to the distance to the nuclear rim (Figure 4.4b) confirming recent observations [Di Primio et al., 2013; Marini et al., 2015; Lelek et al., 2015]. Since LEDGIN treatment shifts integration sites out of transcriptionally active regions comparable to LEDGF/$p75_{KD}$ conditions (Figure 4.3 on page 79), we verified whether LEDGIN treatment might potentially affect the 3D location of integrated provirus as well. Indeed, addition of 3 µM of LEDGIN CX05045 redistributed the integrated provirus towards the inner nuclear compartment. An HIV-1 mutant (HIV-$IN_{A128T/E170G}$), resistant to LEDGINs [Christ et al., 2010], was not redistributed (Figure 4.4d). Therefore, these data demonstrate that LEDGF/p75 controls the nuclear topology of HIV-1 provirus and that LEDGINs may induce its spatial randomization.

4.4.5 The residual reservoir upon LEDGIN treatment is more quiescent.

In a next step, we investigated whether LEDGIN-mediated retargeting also affected the quiescent reservoir, which would be in line with the effects demonstrated earlier in LEDGF/p75 depleted (KD/KO) cells. SupT1 cells were infected with the single-round OGH reporter virus in the presence of increasing concentrations of LEDGIN (CX014442). HIV OGH infection was measured 3 days post infection using flow cytometry resulting in the detection of cell populations carrying both productive ($eGFP^+$, $mKO2^+$ cells) and quiescent ($eGFP^-$, $mKO2^+$ cells) provirus (see also Figure 4.1 on page 73). LEDGIN treatment induced a dose-dependent decrease in the % of $eGFP^+$, $mKO2^+$ cells (Figure 4.5a on page 83) as well as the overall $mKO2^+$ cells (Figure 4.5b). However, similar to LEDGF/p75 depletion (see Figure 4.1c), LEDGIN treatment resulted in a relative increase of the quiescent fraction (% $eGFP^-$, $mKO2^+$ cells) / (% $mKO2^+$ cells) *100 (Figure 4.5c). No increase in the quiescent fraction was observed when adding increasing concentrations of the reverse transcriptase inhibitor AZT (data not shown) suggesting the phenotype was not merely due to inhibition of infection. Interestingly, nearly all infected cells were quiescent at 25 µM LEDGIN (CX014442). Several studies reported on the effect of

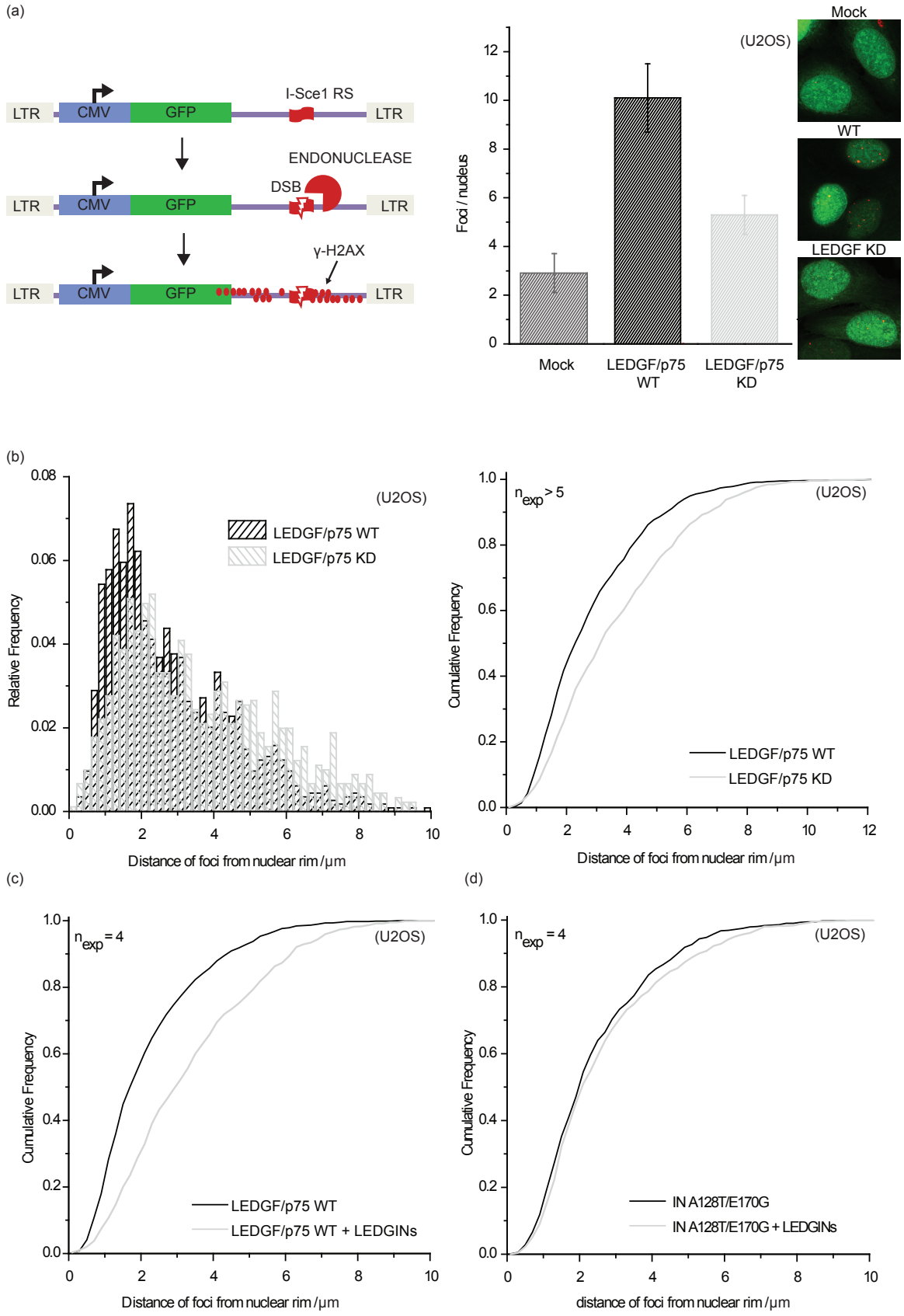


Figure 4.4: LEDGIN treatment shifts HIV-1 proviral localization towards the inner nuclear compartment.

Figure 4.4: LEDGIN treatment shifts HIV-1 proviral localization towards the inner nuclear compartment. U2OS WT or LEDGF/p75 depleted cells were infected with a HIV derived vector pHR-CMV-EGFP with or without an I-SceI restriction site (mock) and γH2AX foci quantified per nucleus 48h post infection after endonuclease digestion. (a) SCIP analysis of proviral DNA corresponding to γH2AX foci (red) in U2OS cells. Bar diagram in the right panel depicts the number of proviruses (γH2AX foci) detected under each condition. Error bars represent standard deviations from at least three experiments. (b) 3D nuclear localization of HIV-1 provirus relative to the nuclear rim in LEDGF/$p75_{KD}$ (empty bars) or WT U2OS cells (grey bars) ($P<0.001$, Kolmogorov-Smirnov test) (n=1000). The right panel depicts the cumulative frequency for the distance relative to the nuclear rim. (c) Cumulative frequency of the 3D nuclear localization of HIV-1 provirus relative to the nuclear rim in U2OS cells treated (grey) with or without (black) 3 µM of LEDGIN CX05045. ($P<0.001$, Kolmogorov-Smirnov test) (n=650). (d) Cumulative frequency of the 3D nuclear localization of HIV-1 provirus relative to the nuclear rim in U2OS cells infected with the HIV-1 $IN_{A128T/E170G}$ and treated with (grey) or without LEDGINs (black, CX05045, 3 µM, 4 x IC_{50}). ($P>0.05$, Kolmogorov-Smirnov test) (n=650). Number of experiments is indicated in each plot (> 100 cells counted/experiment).

integration orientation relative to endogenous genes on the HIV transcriptional state, with a possible enhancement of transcription when integrated in the same orientation or transcriptional interference when integrated in the opposite orientation. We therefore evaluated relative orientation frequencies of those integrations occurring within genes for the different integration site data sets (Supplementary Table S4.3 on page 113). LEDGIN treatment resulted in a significant, dose-dependent increase in the fraction of integrations having an opposite orientation from 46.2 % to 56.7% (p-value < 0.005, Pearsons Chi-square compared to the DMSO control condition) with respect to the targeted gene.

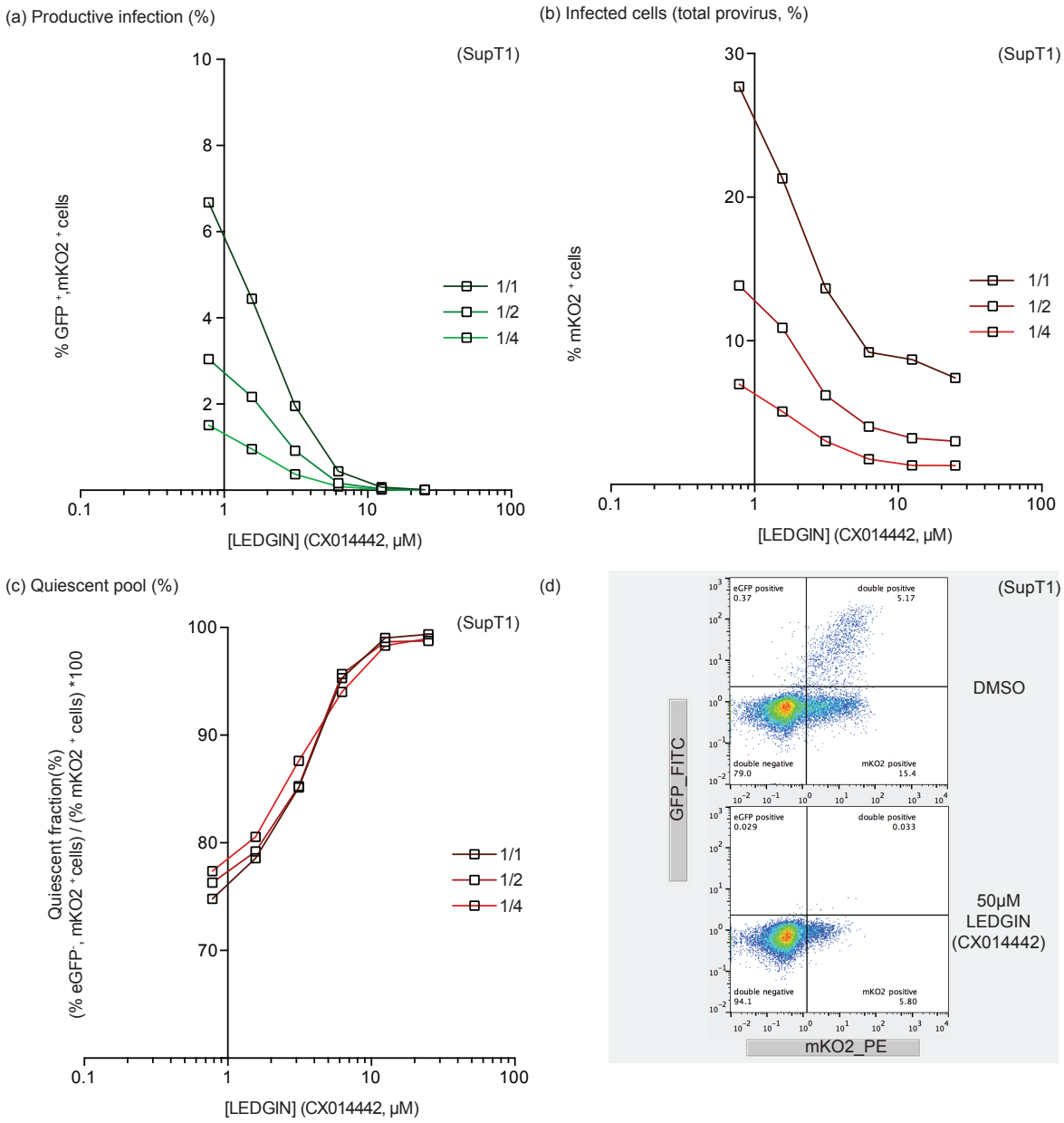


Figure 4.5: LEDGIN mediated retargeting of integration increases the quiescent reservoir. SupT1 cells were infected with three different dilutions of HIV OGH (a) Dose-response curve showing a decrease in % eGFP$^+$, mKO2$^+$ positive cells with increasing LEDGIN concentration. Three different virus concentrations are depicted in green. (b) Dose-response curve showing a decrease in the overall % mKO2$^+$ cells with increasing LEDGIN concentration. Three different virus concentrations are depicted in red. (c) The fraction of quiescent cells (% eGFP$^-$, mKO2$^+$ cells) / (% mKO2$^+$ cells) *100 increases upon addition of LEDGINs. Three different vector dilutions are depicted in red. All viruses are VSV-G pseudotyped. (d) Respresentative dot plots depicting the different cell populations under two different conditions, numbers in the quadrant indicate the percentage of cells. All vectors are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein; mKO2, Mutant Kusubira Orange 2

4.4.6 LEDGIN treatment results in a quiescent reservoir resistant to HIV reactivation.

Next we studied whether LEDGIN treatment also reduces the reactivation potential after reporter gene silencing, as observed under LEDGF/p75 depletion (Figure 4.2 on page 75). The LEDGIN-induced increase in the silent reservoir, together with the reduced HIV reactivation potential could hold promise to reduce the functional reservoir. SupT1 cells were infected with single round HIV-OGH double reporter virus (Figure 4.6 on page 86) or HIV-tCD34 (Supplementary Figure S4.4 on page 108) under varying LEDGIN (CX014442) concentrations and reactivated with LRAs. To demonstrate that LEDGIN-retargeted provirus remains refractory to general cell activation TNFα was used as a reactivation agent. Data depict a representative virus dilution. In Figure 4.6a the percentage of eGFP$^+$, mKO2$^+$ cells and overall % mKO2$^+$ positive cells is plotted after stimulation with DMSO or TNFα for 24 h (open squares and open triangles, for TNFα and DMSO, respectively). Stimulation with TNFα did not affect the percentage living cells (>85% in all conditions 24h post stimulation) neither in the absence or presence of LEDGIN (data not shown). The % eGFP$^+$, mKO2$^+$ cells are significantly higher after stimulation with TNFα compared to DMSO at the different LEDGIN concentrations, while the overall percentages of mKO2$^+$ cells remain constant (compare red and green lines). LEDGIN treatment resulted in reduced reactivation as measured by the fold increase of productively infected cells (% eGFP$^+$, mKO2$^+$) in a concentration-dependent manner, with 25 µM CX014442 reducing the reactivation by TNFα 2-fold (Figure 4.6b). A slope of -0.034 ± 0.009 fold/µM was calculated using a linear regression model (significant deviation from zero; $p<0.0018$, Figure 4.6b). On the other hand, increasing concentrations of LEDGIN CX014442 antagonized the reactivation potential of the latently infected pool as represented by the decrease in the fraction of quiescent cells (% eGFP$^-$, mKO2$^+$ cells) / (% mKO2$^+$ cells)*100 (Figure 4.6c). A slope of -0.93 ± 0.08 %/µM was calculated using a linear regression model (significant deviation from zero; $p<0.0001$, Figure 4.6c). Similar results were observed for HIV-tCD34 where LEDGIN (CX014442) treatment resulted in a dose-dependent inhibition of reactivation from latency (reactivation $IC_{50} \approx 7.24$ µM) as evidenced by the reduced increase in the % tCD34 positive cells (Supplementary Figure S4.4a & b on page 108). This inhibition of reactivation was seen with various LRAs (Supplementary Figure S4.4c). Raltegravir treatment did not result in this phenotype (Supplementary figure S4.4 d & e), excluding integration inhibition as such or an increase in non-integrated 2-LTR circles as the cause of this effect. In conclusion, LEDGIN treatment reduces the reactivation potential of the quiescent HIV pool.

4.4.7 LEDGIN treatment retargets HIV integration into safer locations

The strategy to push HIV into latency by retargeting provirus integration away from active transcription units due to uncoupling of the LEDGF/p75-IN interaction may be associated with an altered risk of insertional mutagenesis. Insertional mutagenesis due to dysregulation of neighboring gene expression is a concern in gene therapy applications with lentiviral vectors [Cavazzana-Calvo et al., 2010]. In addition,

two studies revealed the existence of clonally expanded $CD4^+$ cell populations carrying integrated HIV provirus in HIV-1 patients on prolonged antiretroviral therapy [Maldarelli et al., 2014; Wagner et al., 2014]. Therefore we evaluated whether the residual integration profile under LEDGIN treatment has a different "safety" profile using stringent criteria used in gene therapeutic applications [Papapetrou et al., 2011]. We investigated following criteria defining potentially unsafe integration events: integration near transcription start sites (<50 kb), oncogenes (<300 kb) or miRNA coding regions (<300 kb) and integration into transcription units and ultraconserved elements (UCR). Integration events occurring outside these features are considered to be safe (Table 4.2 on page 88) [Papapetrou et al., 2011]. For each data set, we calculated the percentage of potentially unsafe integration sites according to a given criterion (Table 4.2) and determined the final percentage of safe sites (positioning outside these regions). The most pronounced and dose dependent decrease was observed for integrations falling within transcription units. For the calculation of the final % safe-sites all 5 criteria were taken into account and the risk analysis is therefore affected by the different parameters. In the parental SupT1 cell line only 10.7 % of all vector integration sites can be considered safe in comparison to 28.5% for the matched random control datasets (MRC). LEDGF/p75 depletion increased the percentage safe sites to 14.7% (p-value <0.005, Pearsons Chi-square compared to the DMSO control condition), a phenotype that is reverted upon LEDGF/p75 back complementation (data not shown). The shift in integration site distribution under LEDGIN treatment (Figure 4.3 on page 79 and Supplementary Figure S4.3a-c on page 105), coincides with a dose-dependent increase in percentage of safe integrations (16.51 % at 50 µM, p-value < 0.005, Pearsons Chi-square compared to the DMSO control condition), consistent with the data obtained under LEDGF/p75 depletion.

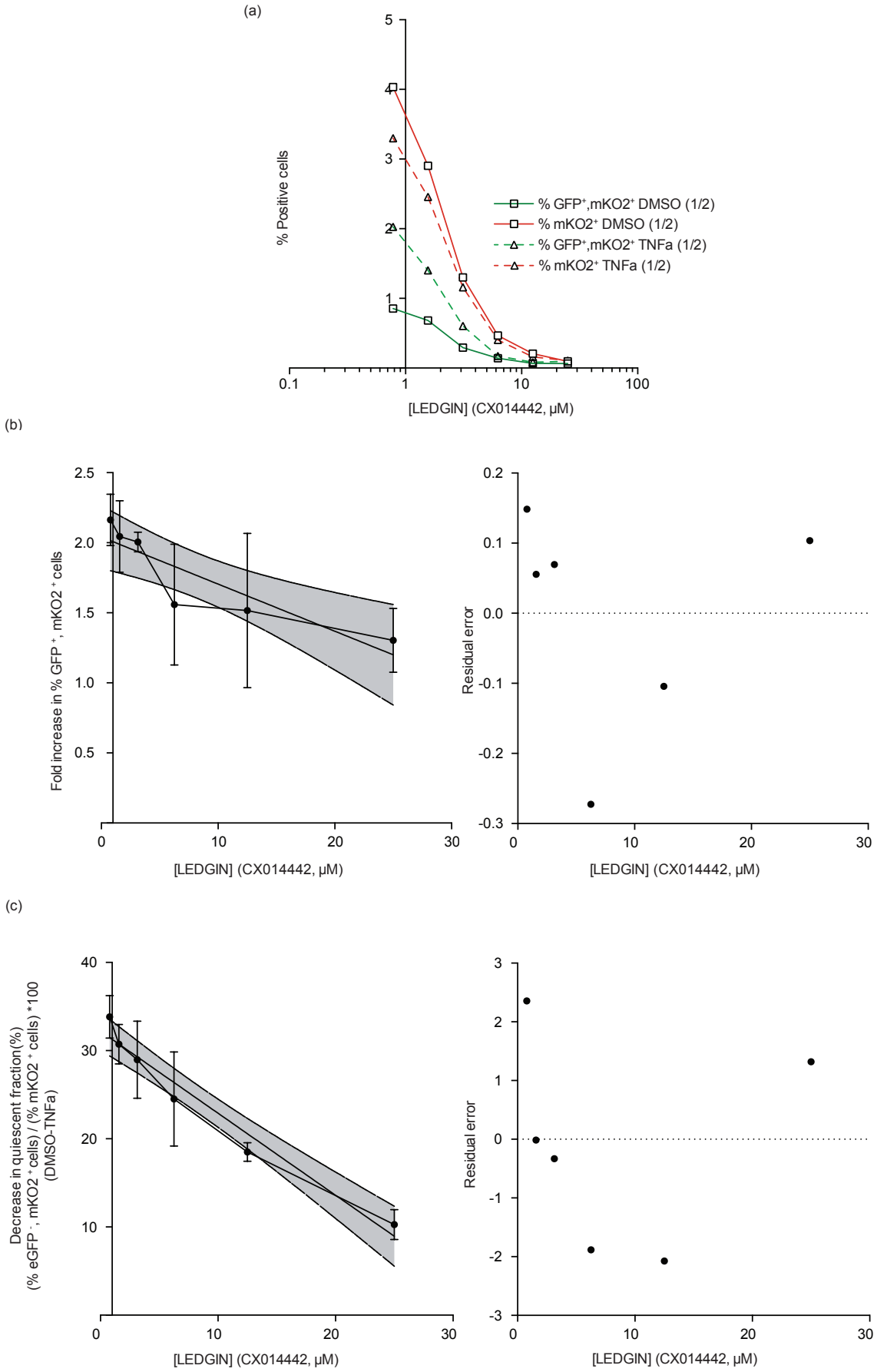


Figure 4.6: LEDGIN treatment reduces reactivation from latency.

Figure 4.6: LEDGIN treatment reduces reactivation from latency. SupT1 cells were infected with single round double reporter virus (OGH) and treated with different concentrations of CX014442 (as indicated). % eGFP - % mKO2 positive cells were monitored, respectively. 11 days post infection cells were reactivated using TNFα (10 ng/mL). (a) Dose-response curve showing the % $eGFP^+$, $mKO2^+$ cells and the overall % $mKO2^+$ cells after reactivation with DMSO or TNFα. (b) Average fold increase in percentage $eGFP^+$, $mKO2^+$ or productively infected cells upon stimulation with TNFα relative to the DMSO treated condition. (c) Average decrease in the fraction quiescent cells (% $eGFP^-$, $mKO2^+$ cells) / (% $mKO2^+$ cells) *100 or silent reservoir fraction upon stimulation with TNFα relative to the DMSO treated condition. Data in (b) and (c) represent averages of 3 different vector dilutions and error bars indicate the standard deviation. The straight lines represent the linear regression calculations together with the 95% confidence band plotted in grey and the residual error relative to the linear regression plot is depicted in the right panel. TNFα; Tumor Necrosis Factor alpha, DMSO; DiMethyl SulfOxide, eGFP; Enhanced Green Fluorescent Protein, mKO2; Mutant Kusubira Orange 2, All viruses are VSV-G pseudotyped.

(SupT1)

Compound concentration (μM)	Type	total.sites	% within 50kb of TSS	% within 300kb of onco	% within 300kb of miRNAs	% in transcript. Units	% in UCR	Total % safe
DMSO Ctrl	insertion	3312	32.04	37.71	25.88	74.68	7.29	10.24
0,78125 (CX014442)	insertion	2451	31.13 *	38.15	25.54	74.95	6.36 *	11.14
1,5625 (CX014442)	insertion	2278	29.28 **	35.47 *	22.52 ***	75.11	7.24	10.36
3,125 (CX014442)	insertion	2485	28.57 ***	35.45 *	23.34 **	71.91 **	7.77 *	11.31 *
6,25 (CX014442)	insertion	3364	29.46 **	33.03 ***	21.76 ***	70.48 ***	6.03 *	13.32 ***
12,5 (CX014442)	insertion	884	31.33	33.03 **	24.89 *	67.42 ***	9.39 **	13.24 **
25 (CX014442)	insertion	604	29.80 *	35.26 *	21.52 **	66.06 ***	6.95	14.57 **
50 (CX014442)	insertion	418	32.54	35.65	22.25 *	59.57 ***	5.98 *	16.51 ***
LEDGF/p75 KD	insertion	4664	32.22	36.16 *	21.67 ***	64.96 ***	6.72 *	14.71 ***
DMSO Ctrl	match	9914	24.16	23.36	14.90	45.07	5.21	29.39
0,78125 (CX014442)	match	7340	24.51	24.55	14.93	45.52	5.38	28.46
1,5625 (CX014442)	match	6828	25.47	23.68	15.39	46.16	5.40	28.31
3,125 (CX014442)	match	7433	23.57	23.92	15.15	46.12	4.94	28.90
6,25 (CX014442)	match	10079	23.96	23.56	15.51	47.10	5.03	28.52
12,5 (CX014442)	match	2649	24.31	24.73	14.65	44.62	5.21	29.26
25 (CX014442)	match	1805	23.32	21.99	15.35	44.76	5.32	29.42
50 (CX014442)	match	1251	25.90	24.62	15.11	46.92	5.28	26.62
LEDGF/p75 KD	match	13961	24.25	24.12	14.83	46.32	5.41	28.66

Table 4.2: Integration frequency near safe harbor criteria. Table showing the percentage of HIV-based vector integration sites relative to features used to define unsafe harbors. These criteria are considered to be unsafe: TSS, Oncogenes, miRNA encoding regions, Transcription units and ultra-conserved regions. The % integrations negatively associated with these 5 features is used to calculate a safety profile. Colors depicts the nature of the association based on the highest and lowest value in the column. Asterisks depict a significant deviation from the DMSO treated control dataset (Pearsons Chi-square test; ***, p-values <0.005; **, p-values <0.05; *, p-values <0.5).TSS, Transcription Start Sites; UCR, Ultra Conserved Regions.

4.4.8 LEDGIN treatment inhibits integration, relatively increases the quiescent viral reservoir and reduces reactivation in primary CD4+ T cells.

Since recent studies reported the existence of a latently infected cell population after infection of activated primary $CD4^+$ T cells [Calvanese et al., 2013; Dahabieh et al., 2013; Chavez et al., 2015], we tried to corroborate the effect of LEDGIN treatment on proviral latency in this model. Human PBMCs were purified, selectively enriched for $CD4^+$ T cells using Bi-specific MAb CD3.8 and infected with the OGH reporter virus, in the presence or absence of LEDGINs (Figure 4.7 a and b on page 90). Similar to the results observed in SupT1 cells, LEDGIN (CX014442) treatment induced a dose-dependent decrease in the % infected cells (decrease in %$eGFP^+/mKO2^+$ cells or overall % $mKO2^+$ cells, Figure 4.7a) and an increase in the fraction quiescent cells (% $eGFP^-$, $mKO2^+$ cells) / (% $mKO2^+$ cells) *100 (Figure 4.7b) reaching 46.3 % quiescence at a CX014442 concentration of 25 µM (representative data from one donor are shown for 2 different donors tested). Next we evaluated the multimodal effect of LEDGIN treatment on integration, assembly and reactivation in a multiple round reactivation model using WT HIV in resting $CD4^+$ T-cells in order to model the in vivo situation. PHA/IL-2 activated primary (resting) $CD4^+$ T cells were infected with NL4.3 virus in the presence of submicromolar concentrations of LEDGINs. At day four post infection (p.i.) LEDGINs were removed and cells were reseeded in the presence of PMA and PHA. Virus production upon reactivation was measured at day 7 p.i. by p24 ELISA. LEDGIN treatment reduced the number of proviral DNA copies in $CD4^+$ T-cells in a dose-dependent manner (Figure 4.7c). These residual integrants were less susceptible to reactivation as displayed by the reduced p24 production (Figure 4.7d, data show the average for two different donors tested). Apart from reducing overall integration, these data suggested that LEDGIN treatment during HIV infection leads to quiescence of the residual integrants both in SupT1 and primary $CD4^+$ T-cells. This quiescent reservoir appears less susceptible to reactivation.

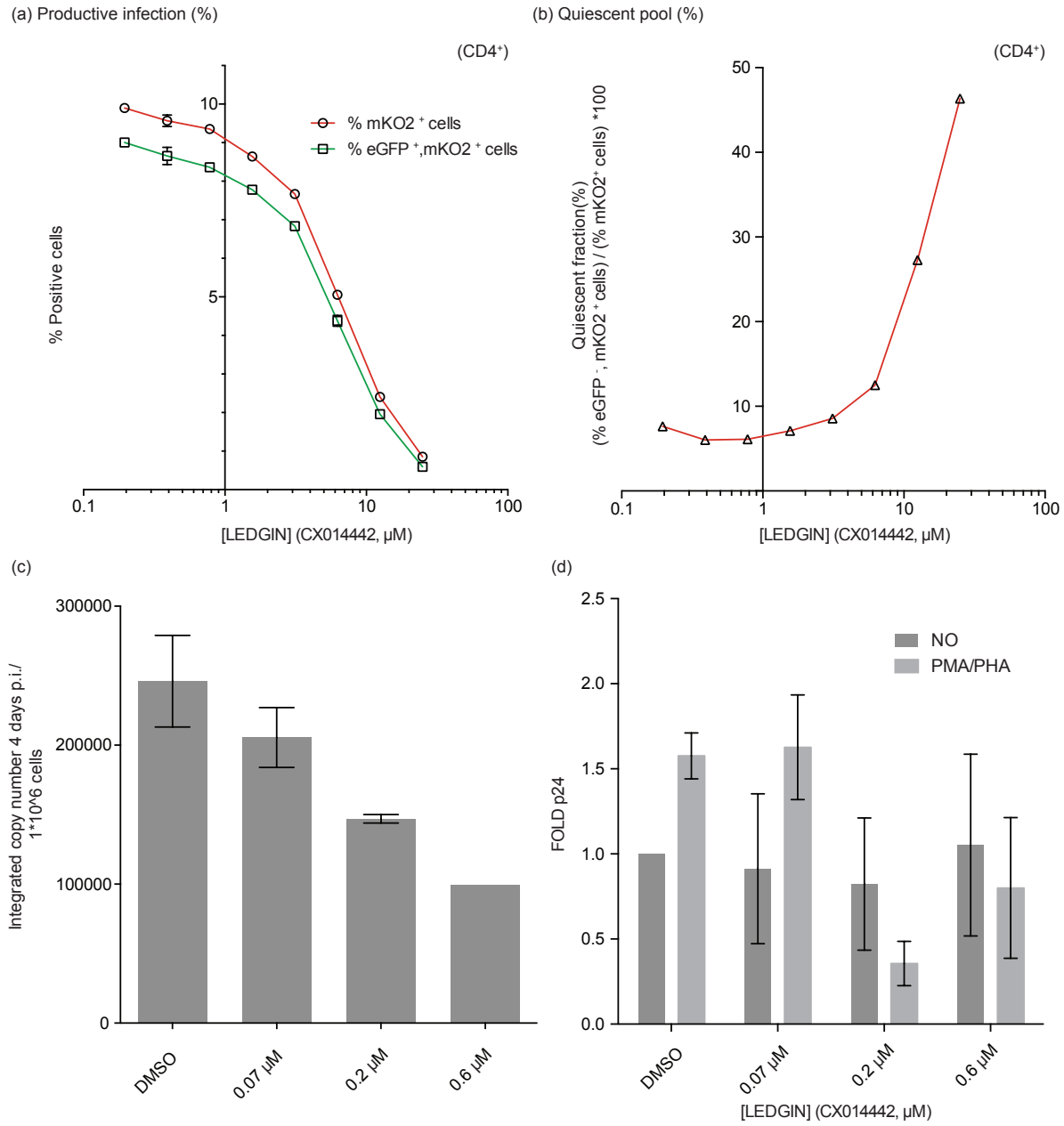


Figure 4.7: LEDGIN treatment inhibits integration, induces quiescence of the residual viral reservoir and reduces reactivation in primary CD4$^+$ T cells. (a) Activated CD4$^+$ T- cells were infected with single round double reporter virus (OGH) and the % eGFP - % mKO2 positive cells were monitored. Dose-response curve shows a decrease both in the % eGFP$^+$, mKO2$^+$ cells and overall % mKO2$^+$ cells with increasing LEDGIN (CX014442) concentration. (b) The fraction of silently infected cell population (% eGFP$^-$, mKO2$^+$ cells) / (% mKO2$^+$ cells) *100 increases upon addition of LEDGINs. Data are representative for two different donors. All vectors are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein; mKO2, monomeric Kusubira Orange 2. (c) Activated CD4$^+$ T- cells were infected with NL4.3 virus under different LEDGIN concentrations. 4 days p.i. integrated copy numbers were determined using a quantitative Alu-LTR PCR. (d) 4 days p.i. CD4$^+$ T-cells were reactivated using PMA/PHA and p24 production in the supernatant was monitored 7 days p.i. by ELISA. Data show the average for two different donors tested ± SEM. PHA, PhytoHaemAgglutinin; PMA, Phorbol 12-Myristate 13-Acetate.

4.4.9 Discussion

The moral duty to respond to the call for an HIV cure calls for exploration of experimental and innovative HIV cure strategies. As a complement to current shock-and-kill approaches aimed at forcing HIV out of its hiding places to obtain a sterilizing cure, we here provide experimental evidence for a strategy to push the virus towards transcriptional quiescence by interference with LEDGF/p75, the main determinant of integration site selection. General belief states that latency is an accident rather than a default pathway of an actively replicating cytopathic virus. The fact that LEDGF/p75, the tethering determinant of HIV integration, directs integration preferentially towards actively transcribed regions, ensuring productive infection, is consistent with this notion. Recently, it was proposed that HIV latency is a hardwired, evolutionarily conserved switch increasing the likelihood of successful mucosal transmission during primary infection [Rouzine et al., 2015; Razooky et al., 2015]. Whereas HIV Tat is known to be the main viral determinant in controlling the HIV transcriptional state, the contribution of other viral and host determinants to the transcriptional state of the HIV provirus awaits further clarification. Here we investigated the role of LEDGF/p75 in the establishment of HIV latency. We demonstrate that LEDGF/p75 depletion, known to result in retargeting of integration away from the body of actively transcribed genes [Schrijvers et al., 2012a; Shun et al., 2007; Fadel et al., 2014], increases the fraction of quiescently infected cells and simultaneously decreases the reactivation potential of the proviruses. In light of the recently proposed role of LEDGF/p75 (and the Iws1/Spt6 complex) in post-integration HIV transcriptional repression [Gérard et al., 2015], this observed decrease may even be an underestimation of the contribution of the integration environment to the latent phenotype. Taken together, these observations suggest that disruption of the LEDGF/p75-IN interaction not only inhibits integration but could as well affect the establishment of the latent pool and the reactivation from latency. Recently developed LEDGINs, *bona fide* small molecule inhibitors of LEDGF/p75-IN interaction, allowed us to test this hypothesis. Indeed here we demonstrate: (i) a LEDGIN-mediated shift in lentiviral integration site distribution resembling LEDGF/p75 depletion (out of the body of actively transcribed genes, with increased integration in the vicinity of CpG islands); ii) a strongly reduced but authentic residual integration, (iii) a LEDGIN-mediated shift in 3D nuclear location of HIV provirus away from the nuclear rim; (iv) a relative increase in the fraction of quiescent proviruses and (v) a dose-dependent block in HIV reactivation from latency both in cell lines and primary $CD4^+$ T-cells. It was recently proposed that pushing enough proviruses into quiescence could drive the basic reproduction number of HIV below 1, resulting in unsustainable infection [Rouzine et al., 2015]. LEDGIN treatment apparently succeeds in rendering (almost) 100 % of the virus into a quiescent state refractory to reactivation (Figure 4.5c on page 83 and Figure 4.6 on page 86). Although final proof will only be obtained in clinical trials, our cell culture data in relevant cell lines and primary cells provide evidence for the feasibility of this strategy. The importance of the site of integration in the human genome for basal transcriptional activity of HIV was evidenced more than a decade ago [Jordan et al., 2001]. Now it is known that genomic target site selection during lentiviral integration is a multi-step process where biases are introduced at different levels, which each in part affect the stochastic proviral gene

expression levels. At least three levels can be recognized. First, nuclear topology and proximity to the nucleopore affect integration site selection [Di Primio et al., 2013; Marini et al., 2015]. Next, chromatin readers such as LEDGF/p75 or HRP-2 that recognize epigenetic marks associated with transcriptional activity tether the preintegration complex to active gene regions. No data exist suggesting a protein gradient. Therefore, as proposed by Marini et al. [Marini et al., 2015], preferential lentiviral integration within the proximity of the nuclear periphery probably reflects the encounter by the HIV PIC of the first LEDGF/p75 bound chromatin close to nucleopores. Finally, bias for target DNA base recognition by integrases also influences local and global integration patterns [Demeulemeester et al., 2014b; Serrao et al., 2015; Demeulemeester et al., 2015]. In theory, interference with any of those mechanisms could shift the resulting proviral reservoir from transcriptionally active to quiescent. Both LEDGF/p75 depletion and LEDGIN treatment affect integration, integration site selection (Figure 4.3 on page 79 and Supplementary Figure S4.3a-c on page 105) and nuclear location (Figure 4.4 on page 81). Probably the reduced reactivation due to altered chromatin context is not entirely reflected by the features measured in the integration site analysis. As shown, aberrant integrations or LTR deletions are most likely not contributing, or only to a minor extent (Supplementary Figure S4.5 on page 110 and Table S4.2 on page 113). In the absence of LEDGF/p75, HRP-2 determines HIV integration site selection. The fact that LEDGINs also inhibit the interaction between HRP-2 and HIV-1 IN [Schrijvers et al., 2012a], can explain the more pronounced effect of LEDGINs on reactivation from quiescence compared to LEDGF/p75-depletion. Identification of the exact nature of the altered chromatin context responsible for the observed phenotype awaits further experimentation. HIV latency is of multifactorial nature and the transcriptional state of integrated provirus is not only influenced by molecular determinants but also depends on the infected host cell and its activation state [Dahabieh et al., 2015]. Therefore the respective quiescent fraction and responsiveness to different LRAs might alter depending on the cell model used since promoter activity may differ between cell types. In order to study the effect of integration site distribution on HIV latency we used NL4.3-based (HIV-tCD34) and LAI-based (OGH) single round reporter viruses in a concise reactivation setup in T-cell lines and activated $CD4^+$ T-cells. LEDGINs block HIV replication during integration (referred to as 'the early effect') as well as during assembly (i.e. 'the late effect'), with inhibition during the late step having a 10- to 100 fold higher potency than the early step [Debyser et al., 2015]. In the experiments with single round virus, we used micromolar concentrations of LEDGIN, required to inhibit the early step. Interestingly, recent experimental results suggest that addition of submicromolar concentrations of LEDGINs during virus production results in viruses that after integration are again refractory to LTR-driven gene expression (Supplementary Figure S4.7 on page 112), suggesting that during multiple round replication also low concentrations of LEDGINs induce quiescent proviral pools. Our data were corroborated with wt NL4.3 virus in IL-2/PHA activated primary $CD4^+$ T-cells (Figure 4.7 on page 90). These multiple round experiments allowed the use of submicromolar LEDGIN (CX014442) concentrations. It will be of interest to extend these observations in the future to other latency models. Preliminary data in a more sophisticated model for HIV latency based on infection of CCL19-activated resting $CD4^+$cells [Spina et al., 2013] corroborates a LEDGIN-mediated shift of the HIV reservoir into a quiescent state

refractory to reactivation (unpublished data). Evidence is growing that initiation of cART early after infection is effective in reducing the size of the viral reservoir [Ananworanich et al., 2012; Hoen et al., 2007; Hocqueloux et al., 2013; Ananworanich et al., 2015; Malatinkova et al., 2015]. Early treatment initiation with ART will likely become standard clinical practice in HIV care. This is supported by the recent outcome of the first large-scale international "Strategic Timing of AntiRetroviral Treatment" (START) study, showing a considerably lower risk of developing AIDS and other serious conditions when compared to later treatment initiation (INSIGHT START Study Group, 2015). We here propose that LEDGINs, currently in (pre-)clinical development as antivirals, could eventually and in synergy with existing antiretrovirals contribute to an HIV remission by acting as potent antivirals with the additional capacity to affect the transcriptional state of the residual HIV reservoir. In an ideal format LEDGINs should be be evaluated in combination with cART regimens initiated during acute infection before reservoirs are established. In light of recent findings on ongoing HIV replication, penetration of drugs in lymph nodes will be required. For HIV pre-exposure prophylaxis (PrEP) LEDGINs have the added benefit that any residual provirus may turn out to be refractory to reactivation. It is clear that this final outcome utterly depends on well-designed clinical trials and surpasses any claims made in this work. Still, we provide initial evidence in cell culture and primary cells for an important role of LEDGF/p75 in the establishment of the replicating reservoir. LEDGINs will allow us to investigate in clinical trials a novel strategy based on a retargeted, more quiescent proviral reservoir. LEDGIN treatment during acute HIV infection may result in a proviruses that are refractory to reactivation after treatment interruption leading to an HIV remission.

4.5 Experimental procedures

4.5.1 Cell culture, virus production and transduction.

All cells were tested to be mycoplasma free. Cells were cultured in a humidified atmosphere containing 5% CO_2 at 37°C. SupT1 (provided by the National Institutes of Health (NIH) Reagent Program, NIH, Bethesda, MD). Nalm cells obtained from ATCC [Schrijvers et al., 2012a] were cultured in RPMI medium (GIBCO-BRL) supplemented with 10% v/v heat inactivated fetal calf serum (FCS, Sigma-Aldrich) and 0.01% v/v gentamicin (GIBCO). HEK293T-cells (gift from O. Danos, Evry, France) were cultured in Dulbecco Modified Eagle Medium (DMEM, GIBCO) with 5% v/v FCS (Sigma-Aldrich) and 0.01% v/v gentamicin (GIBCO). U2OS cells (ATTC) were cultured in DMEM (GIBCO) with 10% v/v FCS. Vesicular stomatitis virus G (VSV-G)-pseudotyped viruses were generated by double transfection of HEK293T-cells with a plasmid encoding a single round HIV clone (pNL4-3.tCD34.R-.E-, pOGH, pOGH-csGFP-only or pOGH-mKO2-only) together with a VSV-G protein encoding plasmid (pVSVG). In other experiments triple transfection was done with the transfer plasmid pHR-CMV-GFP-I-Sce1 together with the δ8.91 packaging plasmid and pVSVG. Linear polyethylenimine (PEI; Polysciences) was used for plasmid transfections. Medium was replaced 6 h post transfection and supernatant collected after 72 h by filtration through a 0.22 µm pore membrane (Corning Inc.). The

virus was concentrated using a Vivaspin 15-50 kDa cut-off column (Vivascience), DNase (Roche) treated and stored at -80°C. The HR vectors were concentrated by 2 h of ultracentrifugation in a 20% sucrose cushion. Cells were seeded and infected for 3 days in 48-well plates (10% FCS, 0.01% gentamicin RPMI) yielding an infection rate < 40% positive cells, as monitored by FACS analysis using a MACS Quant VYB FACS analyzer (Miltenyi Biotech GmbH), ensuring single-copy integration. Cells were washed twice in Phosphate Buffered Saline (PBS) 72 h post infection to remove residual virus and reseeded. FACS samples were taken every 2 days to monitor reporter gene expression. A SIV-based lentiviral vector carrying a spleen focus forming viral (SFFV) promoter driving a zeocin resistance gene and a LEDGF/p75 specific miRNA-based shRNA [Schrijvers et al., 2012a; Osório et al., 2014] was used to generate a SupT1 LEDGF/p75 knock down (KD) cell line. LEDGF/p75 depletion was monitored using Western blot and Q-PCR (>85% in SupT1 LEDGF/p75$_{KD}$ cells, Supplementary Fig. 1a bottom panel). Nalm LEDGF/p75 control (+/c) and LEDGF/p75$_{KO}$ cells (-/-) were generated previously and are described in [Schrijvers et al., 2012a].

4.5.2 Reporter viruses.

Multi-colored reporter virus (OGH). A variant of the recently described LAI-based double reporter virus was used [Chavez et al., 2015], where a constitutive and an LTR-driven reporter are simultaneously measured to study the latent reservoir [Calvanese et al., 2013; Dahabieh et al., 2013]. This orange-green HIV-1 (OGH) reporter virus variant encodes an LTR-driven enhanced Green Fluorescent Protein (eGFP) in the *nef* gene position together with a constitutively active EF1α promoter driving mutant Kusabira-Orange2 (mKO2) expression instead of mCherry as described previously [Calvanese et al., 2013; Chavez et al., 2015] (Figure 4.1a). An internal constitutive promoter driving mKO2 expression allows direct visualization of the LTR-silent latent proviral pool via the FACS measurement of mKO2-based red fluorescence. HIV NL4-3.tCD34.R-.E-. HIV tCD34 is a NL4.3-based single round reporter virus containing the LTR-driven truncated CD34 (tCD34) as a reporter protein in the *nef* gene position (Figure 4.2a). We replaced the firefly luciferase gene in pNL4-3.Luc.R-.E- (NIH aids reagent program) via *NotI-XhoI* digestion with a tCD34 cassette using standard PCR amplification methods. CD34 (cluster of differentiation 34) is a cell surface glycoprotein functioning as a cell-cell adhesion factor in HSCs but is not present on primary CD4$^+$ T lymphocytes [Fehse et al., 2000]. A truncated version was used to block signal transduction and expression was visualized using antibody-staining allowing for non-fluorescent based cell sorting. Human CD34-PE antibody (Miltenyi Biotec, Cat.N°130-081-002) was used to detect tCD34 expression.

4.5.3 Flow cytometry analysis.

Prior to flow cytometry, cells were fixed for 15 min in 4% paraformaldehyde at room temperature. Expression of eGFP/mKO2 or tCD34 was monitored using a MACS Quant VYB FACS analyzer (Miltenyi Biotech GmbH) using a 488 nm, 50 mW DPSS (diode-pumped solid-state) and a 561 nm, 100mW diode laser respectively and 525/50 nm - 586/15 nm band pass filters. A total of at least 30,000 live cells were

counted, as determined on the basis of forward scatter channel/side scatter channel (FSC-H/SSC-H) and doublets were excluded based on the FSC-A/FSC-H or SSC-A/SSC-H plot. Single reporter controls where used for compensation purposes. Data were analyzed using third party software (FlowJo).

4.5.4 Drug treatment.

The LEDGIN CX014442 [Christ et al., 2012] was added at different concentrations during single round infection and washed away together with residual virus 72 h post infection. Samples were harvested for FACS analysis and the remainder of infected cells was reseeded. FACS samples were taken every 2 days to monitor reporter gene expression. The infected cells were reactivated from latency 11 days post infection using Tumor Necrosis Factor alpha (TNFα , 10 ng/mL, Immunosource), suberoylanilide hydroxamic acid (SAHA, 0.3 to 3 µM, AIDS reagents), Prostratin (5 µM, AIDS reagents) or phorbol 12-myristate 13-acetate (PMA, 0.3 µM, AIDS reagents) 24 h prior to analysis by flow cytometry. Time courses and drug concentrations are indicated in the individual experiments. In all conditions the percentage living cells was higher than 75 procent at 24h post LRA administration (except for 3 ÎijM SAHA; 55% living cells) based on FSC/SSC FACS analysis. LEDGINs were synthesized at Cistim/CD3 KU Leuven (courtesy of Dr. A. Marchand).

4.5.5 Genomic DNA isolation and quantification of Integrated copy number.

Two million cells were pelleted and genomic DNA extracted using a mammalian genomic DNA miniprep kit (Sigma-Aldrich). Standard spectrophotometric methods were used to determine the genomic DNA concentration. Samples corresponding to 250 ng genomic DNA were used for analysis. Each reaction contained 12.5 µl iQ Supermix (Biorad), 40 nmol/L forward and reverse primer (5' TGCACCCTGTGTCTCAACAT 3' and 5' GGCTTCAAGGTTGTCTCTGG 3' respectively) and 40 nmol/L of tCD34 probe (5' (6FAM)-GGCCACAACAAACATCACAG-(TAM) 3') in a final volume of 25 µl. In all cases, RNaseP was used as an endogenous control for normalization (TaqMan RNaseP control reagent, Applied Biosystems, The Netherlands). Samples were run in triplicate for 3 minutes at 95 °C followed by 50 cycles of 10 seconds at 95 °C and 30 seconds at 55 °C in a LightCycler 480 (Roche-applied-science). Analysis was performed using the LightCycler 480 software.

4.5.6 Integration site amplification.

Integration sites were determined as described previously [Marshall et al., 2007]. In short, cells were seeded and transduced with a lentiviral vector for 3 days, then washed twice with PBS. Transduced cells were further cultivated for at least 10 days to eliminate non-integrated DNA. Cells were harvested and genomic DNA extracted using the GenElute Mammalian Genomic DNA miniprep kit (Sigma-Aldrich). Integration sites were amplified by linker-mediated PCR as described previously [Marshall et al., 2007] (see Supplementary Figure A.1). Genomic DNA was fragmented using *MseI* restriction

digestion and linkers ligated. Provirus/host genome junctions were amplified by nested PCR using indexed primers. Products were gel-purified and sequenced using 454/Roche pyrosequencing (Titanium technology, Roche) on the 454 GS-FLX-instrument. Reads were filtered based on perfect matching of the LTR linker, barcode and flanking LTR. All sites were mapped to the human reference genome requiring a perfect match within 3 bp of the LTR end. Matched Random Control (MRC) sites were computationally generated and matched to experimental sites with respect to the distance to the nearest MseI cleavage site. Normalization of experimental HIV sites by the MRC sites corrects for the recovery bias due to cleavage by *MseI*. Analysis was performed as described previously [Marshall et al., 2007]. A more detailed guide to the data presented can be found in [Ocwieja et al., 2011]. Sequence logos were created using WebLogo 3.3 with compositional adjustment for the human genome base background distribution.

4.5.7 Single Cell Imaging of Proviral HIV (SCIP) assay.

20,000 U2OS cells/well were seeded in a 4 well chamber slides (Lab-TekTM) and transfected the next day with 200 ng of pCBASce plasmid encoding the I-Sce1 endonuclease using Effectene (Qiagen). Six hours post transfection cells were infected with 4 reverse transcriptase units (RTUs; SYBR Green-based Product Enhanced Reverse Transcriptase assay, SGPERT; [Pizzato et al., 2009] of the lentiviral vectors HR-CMV-GFP-I-SceI/VSV-G or HR CMV-GFP-I-Sce1-INA128T/E170G/VSV-G in Optimem medium containing 1% FCS for 2 h and fixed 48 h post infection with 4.0% paraformaldehyde in PBS for 5 minutes at room temperature (RT). After permeabilization with PBS containing 0.2% Triton-X100 for 10 minutes, samples were blocked overnight with 3% Bovine Serum Albumin (BSA) at 4°C. The slides were incubated with the primary antibody directed against phosphorylated γH2AX (1:500, 05-636 Millipore) for 1 h at RT and with secondary fluorophore-conjugated antibodies for 1 h at RT. Slides were mounted with Vectashield mounting medium (Vector Laboratories). Nuclear fluorescent signal from γH2AX foci was acquired with the TCS SL laser-scanning confocal microscope (Leica Microsystems) equipped with galvanometric stage using a 63x/1.4 NA HCX PL APO oil immersion objective and processed by an image software (ImageJ, NIH.gov) [Di Primio et al., 2013]. Background, i.e. spontaneous γH2AX repair foci, were subtracted before further analysis as described in [Di Primio et al., 2013].

4.5.8 CD4+ T-cell enrichment.

Human peripheral blood mononuclear cells (PBMCs), obtained from the Red Cross Blood transfusion Center (Mechelen, Belgium) according to approved bioethical guidelines of our institute (S57175-IRB00002047) , were purified from fresh buffy coats using lymphoprep density gradient centrifugation (Stem cell technologies). The CD4+ T-cells were selectively enriched using Bi-specific MAb CD3.8 (0.5 µg/mL, AIDS reagents) for 5 days. Cells were cultured in RPMI 1640, 15% v/v FBS, 0.1% v/v Gentamicin, 100 U/ml IL-2 (Peprotech) (T-cell medium, TCM). Enriched total $CD4^+$ primary T-cells were infected with single round reporter virus for 2 h at 37°C, washed twice in TCM and reseeded in

medium containing different concentrations of LEDGIN CX014442. HIV infection was monitored 48h post infection using flow cytometry analysis.

4.5.9 Reactivation of latent provirus in primary CD4+ T-cells.

Peripheral blood mononuclear cells (PBMCs) were isolated from fresh buffy coats obtained from the Red Cross Blood transfusion center (Mechelen, Belgium) according to approved bioethical guidelines of our institute (S57175-IRB00002047). Resting $CD4^+$ T-cells were purified using a custom-made EasySep negative selection kit (Stem Cell Technologies; 19052 cocktail, with the addition of CD25, CD69, and HLA-DR antibodies (catalogue number 19309VK)). The resulting 95% pure resting $CD4^+$ T-cells consisted of both naive and central memory T-cells [Sallusto et al., 1999]. These freshly isolated resting $CD4^+$ T-cells were activated with 10 µg/ml PHA (Sigma-Aldrich) and 100 U/mL IL-2 (Peprotech) for 2 days before infecting with NL4.3 wt virus for 2 hrs (3.5 * 103 ng p24 per 1 * 10E7 cells/mL). Cells were washed twice with PBS and reseeded in the presence of varying concentrations of LEDGIN (CX014442) and 1 U/mL of IL-2. Four days post-infection cells were washed twice using PBS and some cells were harvested for quantification of integrated proviral DNA using real-time PCR (nested Alu-LTR PCR, [Butler et al., 2001; Lewin et al., 2008] normalized for input DNA by qPCR for the CCR5 gene as previously described [Zhang et al., 1999]. Other cells were reseeded in the presence of 10 nM PMA (Sigma-Aldrich) together with 10 µg/mL PHA (Sigma-Aldrich) or left untreated. PHA activated feeder peripheral blood mononuclear cells (PBMCs) were added 24 h after the activating stimulus to amplify virus replication and enhance detection of the infection [Saleh et al., 2011]. Virus production was measured in culture supernatant at day 7 post-infection by p24 ELISA (Fujirebio Europe).

4.5.10 Statistical analysis.

Reactivation results are expressed as means ± standard error of the mean. Statistical analysis was assessed using multiple t-tests and corrected using Sidak-Bonferroni with significance levels indicated. Ranked Wald statistics were used to calculate the statistical significance (asterisks) for a given genomic feature between integration site datasets relative to the DMSO treated condition (dashes). Significant deviation from the DMSO treated control dataset for safe harbor criteria was calculated using a Pearsons Chi-square test. ImageJ software was used to measure the relative distance of γH2AX foci to the nuclear rim. Statistical differences were calculated using a Kolmogorov-Smirnov test as described previously [Di Primio et al., 2013].

4.5.11 VISA-trimming non-genomic portions of sequence reads.

24 bp of the LTR sequence are used to generate the 'LTR substrings' for substring matching. 3 overlapping LTR substrings are generated, each consisting of 12 bp. LTR substring 1 covers 12 bp located 3' of the LTR sequence, LTR substring 2 is shifted 6 bp upstream from the 3' end of the LTR sequence, and LTR substring 3 is shifted 12 bp upstream from the 3' end of the LTR sequence

(Supplementary Figure S4.6). As our primers are binding 9 bp upstream of the LTR end the 2nd and 3rd string will overlap with our primer. LTR substring 1 is examined first. If LTR substring 1 is found, the query sequence begins immediately following its position and the LTR-chromosome junction is deemed 'normal'. If a mismatch in the LTR portion of the sequence read prevents LTR substring 1 to be found, VISA then searches for LTR substring 2. If LTR substring 2 is found, the query will begin 6 bp downstream of its position to compensate for the 6 bp shift. If a mismatch in the LTR portion of the sequence read prevents LTR substring 2 to be found, VISA searches for LTR substring 3. If LTR substring 3 is found, the query will begin 12 bp downstream of its position to compensate for the 12 bp shift (Supplementary Figure S4.6). When either LTR substring 2 or LTR substring 3 is found, the LTR-Chromosome junction is deemed 'imperfect'. If a mismatch causes LTR substring 3 not to be found, the above steps are repeated using the sequence read in its reverse orientation. If none of the LTR substrings are found in either orientation, the sequence read does not contain a valid query and will not be aligned to the genome. A similar approach was used to trim linker cassettes. A more detailed explanation on trimming non-genomic portions of the sequence reads can be found in (Hocum et al., 2015).

4.5.12 Virus production in the presence of LEDGINs

Vesicular stomatitis virus G (VSV-G)-pseudotyped single-round virus particles were generated by double transfection of HEK293T-cells with a plasmid encoding a single-round HIV clone (pOGH) together with a VSV-G protein encoding plasmid (pVSVG). Linear polyethyleneimine (PEI; Polysciences) was used for plasmid transfections. Cells were washed 6 hrs post transfection (3x with PBS) and medium was added supplemented with a dilution series of LEDGIN CX014442 (0.27 µM - 0.00375 µM). Supernatant was collected after 48 hrs by filtration through a 0.45 µm pore membrane (Corning Inc.). The virus was concentrated using a Vivaspin 15-50 kDa cut-off column (Vivascience), washed 3 times thoroughly with PBS in order to remove residual compound, DNase (Roche) treated and stored at -80°C. Virus productions from at least 10 petri dishes were used for each condition to reduce variation in production efficiency. Productions were normalized for p24/RT. SupT1 and THP1 cells were seeded and infected with equal RT units for 2 hours, washed twice and seeded in 48-well plates (10% FCS, 0.01% gentamicin RPMI). Cells were harvested 48 hours post infection yielding an infection rate $< 40\%$ positive cells, as monitored by FACS analysis using a MACS Quant VYB FACS analyzer (Miltenyi Biotech GmbH), ensuring single-copy integration.

4.6 Author contributions

Z.D. conceived the study. L.V., A.B., S.S., G.V., C.W. and R.S., designed and conducted cell culture experiments. L.V., A.B, S.S., G.V., C.W., R.S., A.C., F.C., R.G. and Z.D analyzed the data. J.D and L.V. performed bioinformatic analyses. E.B., E.V, R.G. and R.S. cloned and designed the different reporter viruses. L.V., F.C., R.G. and Z.D. prepared the manuscript. Z.D. and R.G. were responsible

for the coordination of the study. All authors read, corrected, and approved the final manuscript.

4.7 Acknowledgements

We are grateful to S. Saleh and J. De Rijck for critical reading and thank P. Van de Velde, B. Vanremoortel and NJ. Van der Veeken for their technical assistance. Viral vector production was performed at the Leuven Viral Vector Core. LEDGINs were synthesized by Cistim/CD3 (courtesy of Dr. A. Marchand). L. Vranckx is a doctoral fellow supported by the Flemish Fund for Scientific Research (FWO; Fonds voor Wetenschappelijk Onderzoek). Research at KU Leuven received financial support from the FWO, the KU Leuven Research Council (OT; OT/13/098), HIV-ERA EURECA (IWT-SBO-EURECA), the KU Leuven IDO program IDO/12/008) the Belgian IAP Belvir and the Creative and Novel Ideas in HIV Research Program (CNIHR) through a supplement to the University of California at San Francisco (UCSF) Center For AIDS Research funding (ZKC7449, P30 AI027763). This latter funding was made possible by collaborative efforts of the Office of AIDS Research, the National Institute of Allergy and Infectious Diseases, and the International AIDS Society.

4.8 Disclosure statement

F.C. and Z.D. are inventors on LEDGIN patent applications but the specific compounds used in this paper do not fall under those patent applications.

4.9 Supporting information

4.9.1 Supplementary results

4.9.1.1 LEDGIN treatment shifts HIV integration out of transcription units.

Integration site data sets obtained from SupT1 cells infected with LVP2A and treated with different concentrations of CX014442 were compared to different genomic (Supplementary Figure S4.3a on page 105) and epigenetic features (Supplementary Figure S4.3b-c). A heat map was generated using the INSIPID software (Bushman Lab, University of Pennsylvania). Tile color depicting the nature of the correlation for an integration data set with the respective genomic/epigenetic feature (rows, left) relative to matched random controls, as indicated by the colored Receiver Operating Characteristic (ROC) curve area scale at the bottom of the panel. Columns indicate different data sets. Statistical significance (asterisks, ranked Wald tests) is shown relative to the DMSO or LEDGF/p75$_{KD}$ population (dashes) **p<0.01; ***p<0.001. Significance is reached when p<0.001, compared to the DMSO or LEDGF$_{KD}$ respectively control. WT lentiviral integration correlates with histone marks associated with transcriptionally active chromatin (H3K4 mono-, di- and trimethylation, H3K14 and H4 acetylation, as well as acetylation and monomethylation of H3K9/K27/K79, H4K20 and H2BK5, etc.) (Supplementary Figure S4.3b-c) [De Ravin et al., 2014]. In line, integration in regions associated with marks common to

transcriptionally silent regions or heterochromatin was disfavored (H3K27me$_3$, H3K9me$_3$ or H4K20me$_3$ and H3K79, respectively) (Supplementary Figure S4.3b-c)[De Ravin et al., 2014] corroborating the preference for open, transcriptionally active chromatin. CX014442 treatment during infection shows a dose dependent shift out of transcriptionally active regions as supported by both the genomic and epigenetic analysis. (Supplementary Figure S4.3a-c, a more detailed guide to the data presented can be found in [Ocwieja et al., 2011]).

4.9.1.2 LEDGIN treatment induces a silent reservoir resistant to HIV reactivation.

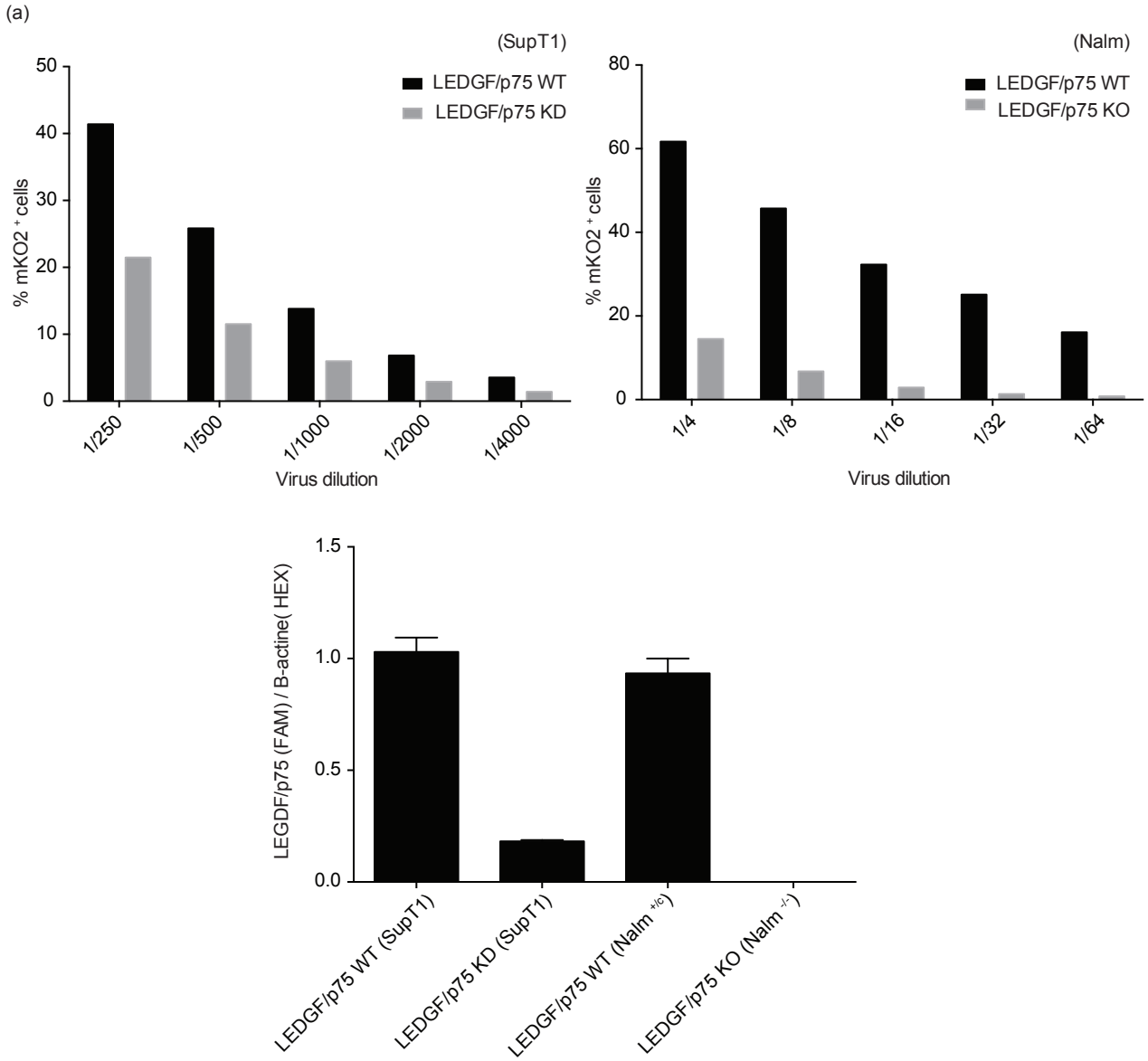
SupT1 cells were infected with single round HIV-tCD34 at a MOI yielding < 30 % positive cells and treated with 0.3-25 µM LEDGIN CX014442. The infected cells were cultured for two weeks in order for silencing to occur and stimulated 11 days post infection using TNFα (10 ng/mL) in order to obtain a maximal T-cell activation response, or DMSO as a control (closed circles and open circles, respectively; Supplementary Figure S4.4a-b on page 108). LEDGIN (CX014442) treatment resulted in a dose-dependent inhibition of reactivation from latency (reactivation $IC_{50} \approx 7.24$ µM) as evidenced by the reduced increase in the % tCD34 positive cells (Supplementary Figure S4.4a & b). The experiment was repeated with a similar reactivation setup. SupT1 cells were infected with single-round HIV-tCD34 during LEDGIN treatment at concentrations equal to 1 and 3 times the IC_{50} of CX014442 (3 and 12 µM respectively, as described above) (Supplementary Figure S4.4c). Cells were reactivated at day 11 with the same agents used in the LEDGF/p75-depleted cells. No difference was observed between the no-stimulation and DMSO control condition. A modest 2-fold tCD34 reactivation was observed after addition of PMA or Prostratin, and an increase in the percentage tCD34 positive cells of 3.5-fold was observed after stimulation with 10 ng/mL TNFα or 3 µM SAHA (Supplementary Figure S4.4c). The percentage living cells amounted at least 75% of the total cell population 24h post stimulation (in the presence of 3 µM SAHA 55% of cells were alive) and was not substantially different upon LEDGIN treatment (data not shown). When comparing reactivation for the LEDGIN-treated conditions to the DMSO control we observed an overall reduction in reactivation potential for the different LRAs (significance was calculated using multiple t-tests and corrected using Sidak-Bonferroni as depicted in the figure (* $p < 0.05$, ** $p < 0.005$, ***$p < 0.0005$)).

4.9.1.3 LEDGIN treatment does not affect the local integration site preferences.

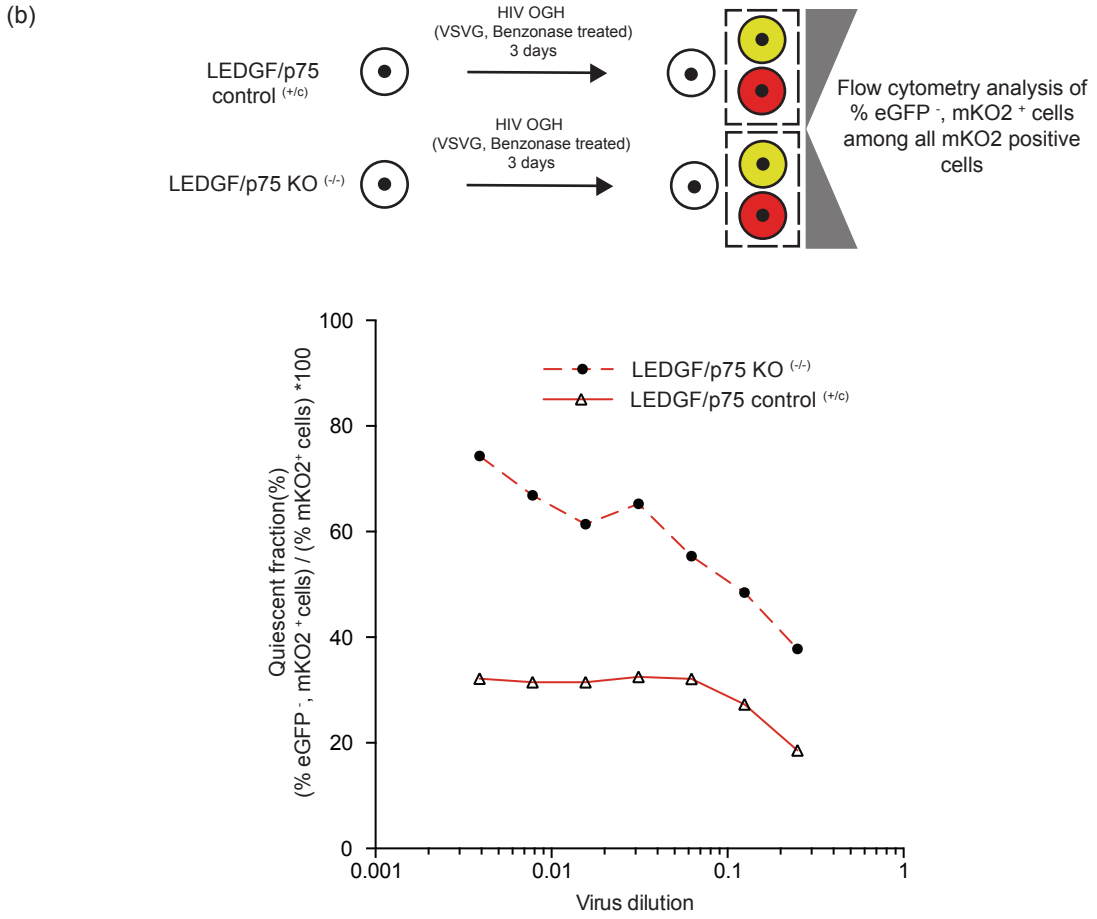
Retroviral INs show a weak but discernable target sequence specificity at the site of integration. LEDGIN-mediated inhibition of the LEDGF/p75-IN interaction by binding to the LEDGF/p75 binding pocket in the catalytic core domain (CCD) dimer interface could potentially affect the inherent integration mechanism, which may result in an increase of aberrant integrations or an altered local integration site preference. We therefore evaluated sequence conservation and relative base frequency in the 18 bp genomic DNA sequence surrounding the integration sites (corresponding to the intasome footprint) using sequence logos (Supplementary Figure S4.5 on page 110). Sequence conservation is indicated as the total height of each stack (measured in bits), while the relative height of bases in a

stack reflects the base frequencies at that position. The site of strand transfer on the plus target DNA strand is depicted as position 0. When comparing pannels a-h, the WT local palindromic sequence logo surrounding the integration site is observed in all conditions. With integration sites being amplified using primers designed to bind 9 bp upstream of the LTR end, we were able to look into sequence variations in those 9 bp and evaluated whether LEDGIN treatment increased the number of aberrant integration events. Virus-host DNA junctions were screened and variations close to the viral-host NA junction evaluated using the VISA platform (Vector Integration Site Analysis server [Hocum et al., 2015]. VISA uses a Perl substring matching strategy to detect and remove the non-genomic sequences from the sequence reads to generate the queries for alignment (Supplementary Figure S4.6 on page 111) [Hocum et al., 2015]. We calculated the percentage unique integration sites with an imperfect LTR chromosome junction relative to the total integration sites for the respective data sets and plotted the different percentages (Supplementary Table S4.2 on page 113). Only a minor fraction (<2 %) of the total integration sites was identified as containing imperfect LTR-chromosome junctions. At 50 µM of LEDGIN a modest increase in the percentage of aberrant junctions was detected. In an effort to characterize this low fraction of aberrant integrations (<2 %) in more detail, we analyzed sequence conservation and relative base frequency. Nucleotide substitutions rather than indel formation, represented by high relative base frequencies, were observed for each of the last 7 bp of the LTR (position -7 to -1) and resembled the genuine LTR sequence end (TCTAGCA) (data not shown). The result was identical in WT conditions and in LEDGIN treatment conditions. Even though this is a single measurement, the observed percentages cannot explain the phenotype observed whereby LEDGINs induce a residual proviral pool that is refractory to reactivation. This analysis allows us to conclude that even though LEDGINs retarget lentiviral integration towards more random, the actual residual insertion event is catalyzed by the viral integrase, which dictates the local integration site preference much alike the WT intasome.

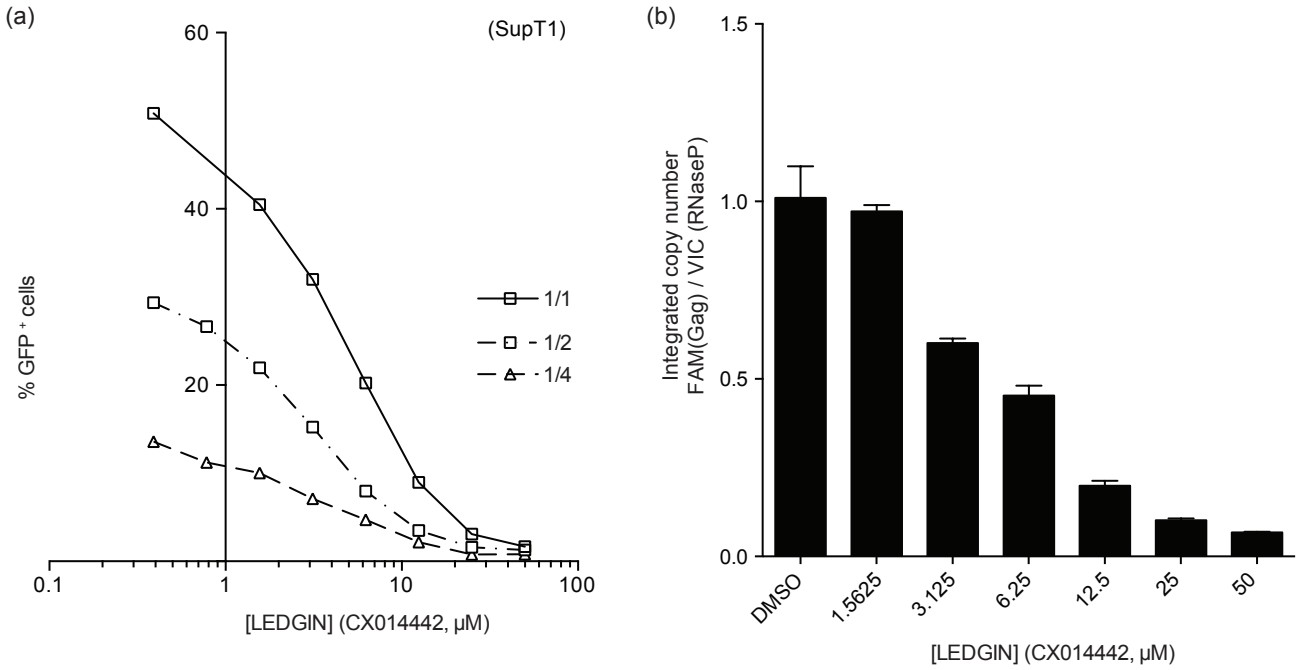
4.9.2 Supplementary figures



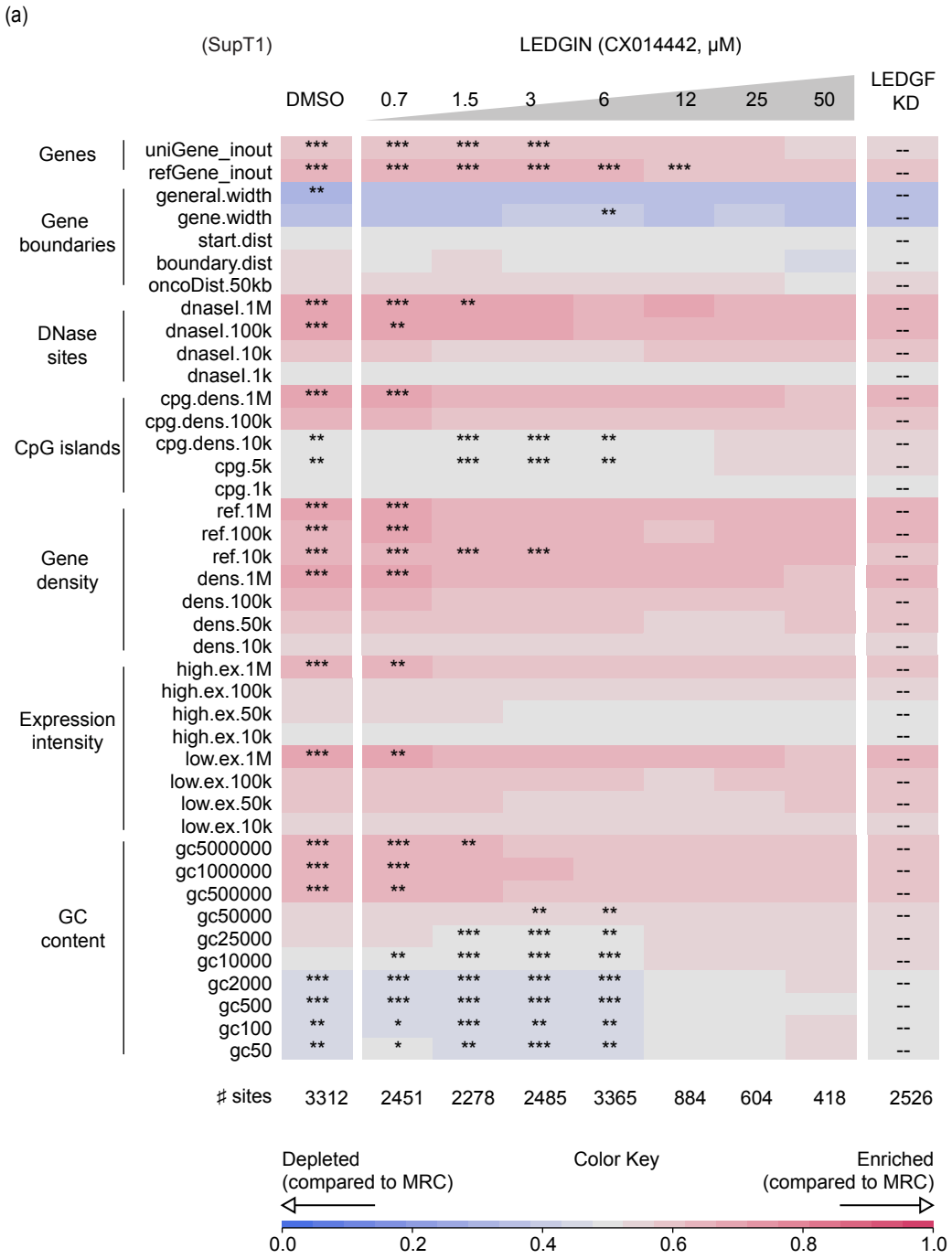
Supplementary Figure S4.1: LEDGF/p75 depletion relatively increases the silent reservoir. (a) LEDGF/p75$_{KD}$ reduces transduction efficiency in LEDGF/p75 depleted cells. Left panel depicting the overall % mKO2 positive cells for SupT1 LEDGF/p75$_{WT}$ and KD cells when infected using different concentrations of OGH (VSV-G). Right panel depicting the overall % mKO2 positive cells for Nalm control (+/c) and LEDGF/p75$_{KO}$ (-/-) cells when infected using different concentrations of OGH (VSV-G). Bottom panel depicting the LEDGF/p75 mRNA levels relative to β-actin levels for de respective cell lines. All viruses are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein; mKO2, Mutant Kusubira Orange 2.



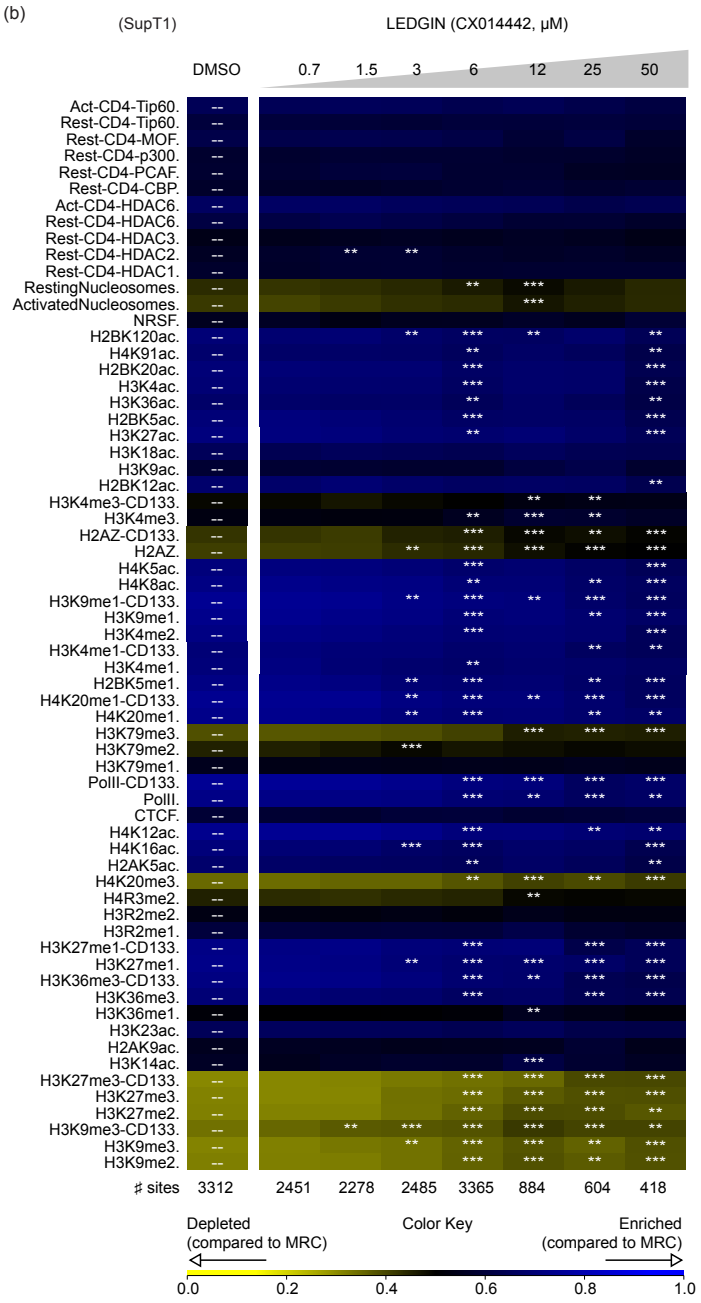
Supplementary Figure S4.1: LEDGF/p75 depletion relatively increases the silent reservoir. (b) LEDGF/p75$_{KO}$ increases the fraction of silently infected cells (% eGFP^{-}, mKO2^{+} cells / % mKO2^{+} cells *100) in Nalm LEDGF/p75$_{KO}$ cells (-/-) compared to Nalm LEDGF/p75 control (+/c). All viruses are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein; mKO2, Mutant Kusubira Orange 2.



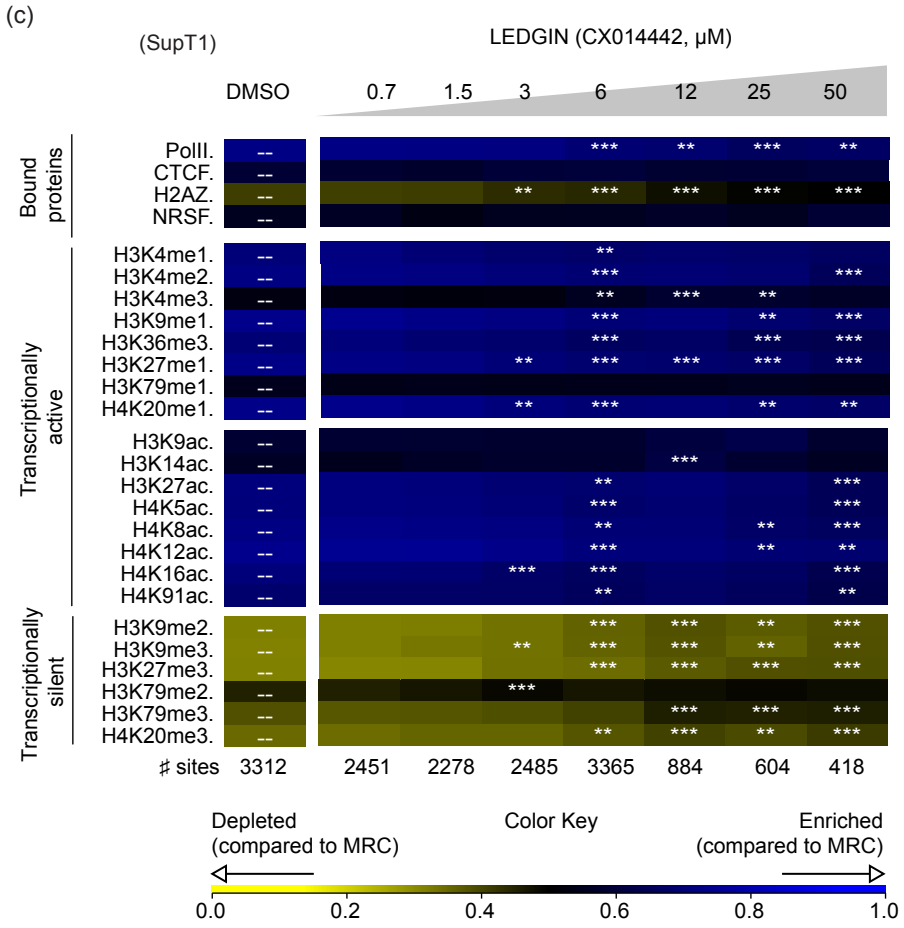
Supplementary Figure S4.2: LEDGIN-mediated inhibition of lentiviral transduction. SupT1 cells were infected with three different dilutions of a lentiviral vector expressing both eGFP and the firefly luciferase (LV eGFP-fluc). (a) The dose-response curve shows a decrease in % eGFP positive cells with increasing LEDGIN concentration. (b) Panel depicting a decrease in integrated copies as determined by Q-PCR analysis. (Gag normalized to RNaseP, data represent the average of 3 measurements ± SD). All vectors are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein.



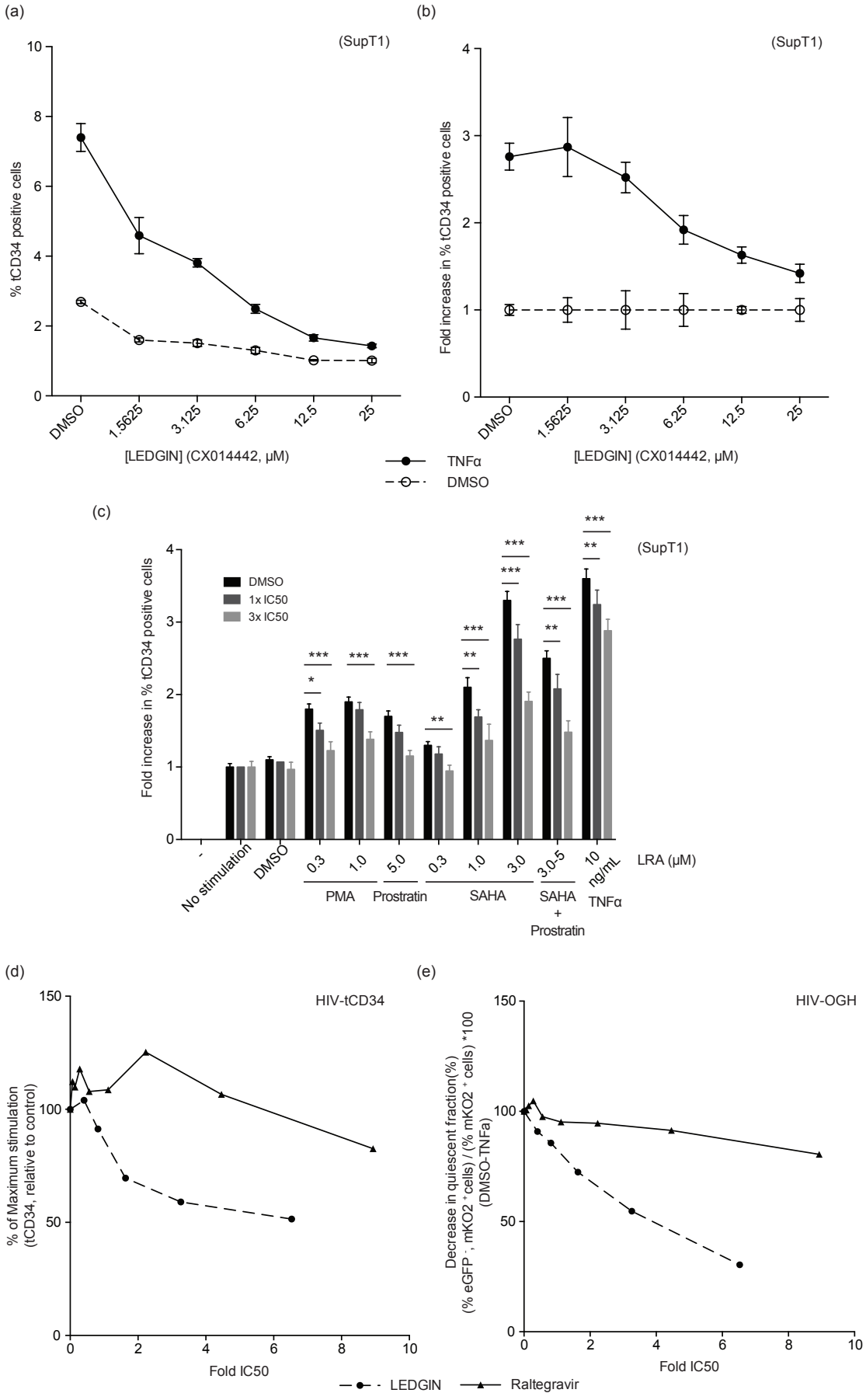
Supplementary Figure S4.3: LEDGIN inhibition of the LEDGF/p75-IN interaction retargets lentiviral integration. Integration site data sets obtained from SupT1 cells infected with LV eGFP-fluc (LVP2A) and treated with different concentrations of CX014442 were compared to genomic and epigenetic features. Heat maps were generated using the INSIPID software (Bushman Lab, University of Pennsylvania). Tile color depicts the nature of the correlation for an integration dataset with the respective genomic/epigenetic feature (rows, left) relative to matched random controls, as indicated by the colored receiver operating characteristic (ROC) curve area scale at the bottom of the panel. Columns indicate different data sets analyzed. (a) Genomic heat map comparing different genomic features (windows are indicated, a more detailed guide to the data presented can be found in [Ocwieja et al., 2011]). Statistical significance (asterisks, ranked Wald tests) is shown relative to the LEDGF/p75$_{KD}$ population (dashes) **p<0.01; ***p<0.001. Significance is reached when p<0.001, compared to the LEDGF/p75$_{KD}$ control.



Supplementary Figure S4.3: LEDGIN inhibition of the LEDGF/p75-IN interaction retargets lentiviral integration. (b) Epigenetic heat map comparing different epigenetic feature (10kb windows). Statistical significance (asterisks, ranked Wald tests) is shown relative to DMSO population (dashes) **p<0.01; ***p<0.001. Significance is reached when p<0.001, compared to MRC. Detailed information on epigenetic marks and their roles can be found in [Barski et al., 2007; Taverna et al., 2007]. Included features were limited to those identified in high-throughput studies in HeLaP4 and primary $CD4^+$ T-cells. CX014442 treatment during infection induces a dose dependent shift out of transcriptionally active regions. Lower significance is observed at the highest LEDGIN concentrations due to a lower copy number of integration sites. DMSO, Dimethyl sulfoxide.

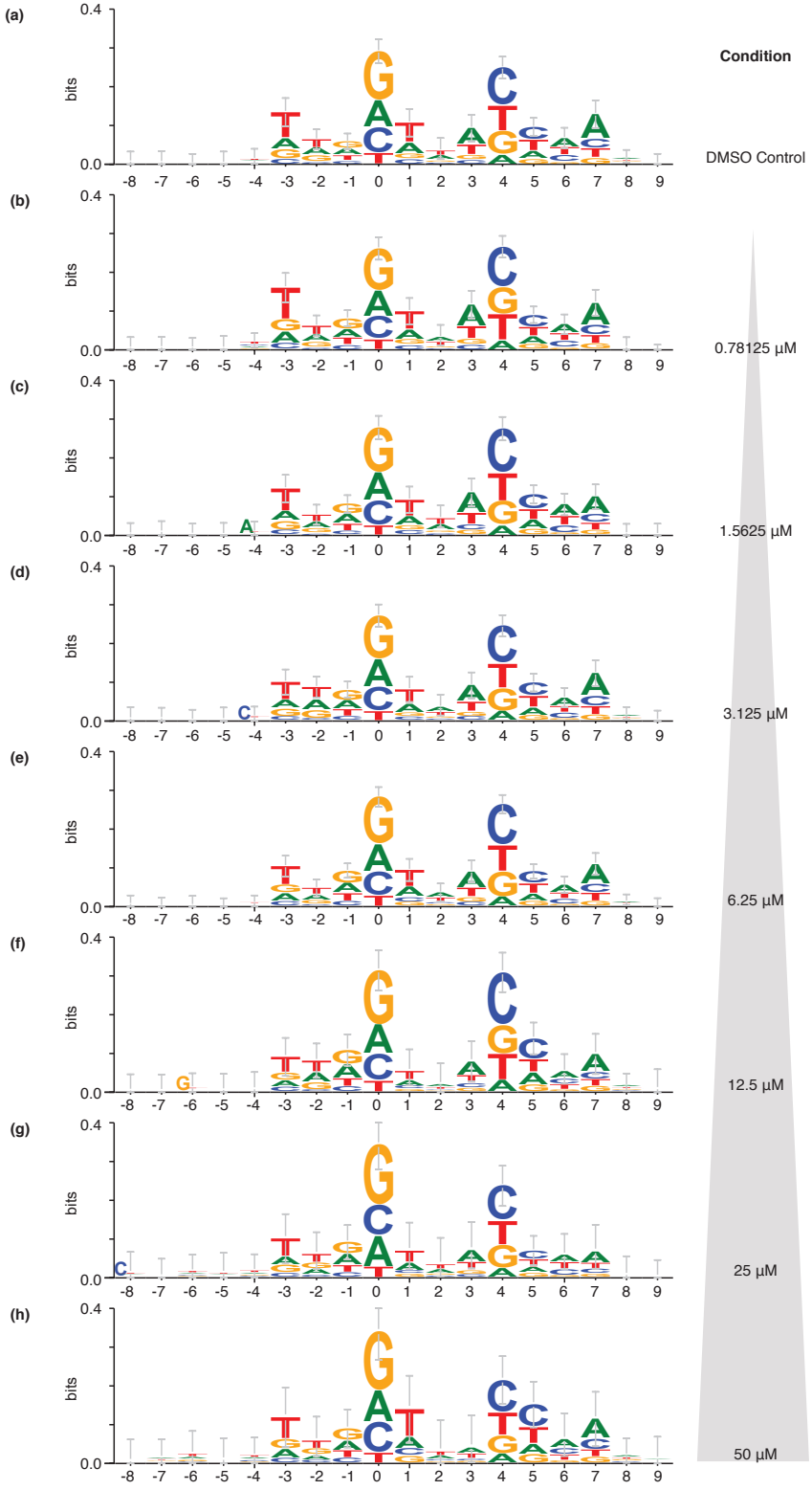


Supplementary Figure S4.3: LEDGIN inhibition of the LEDGF/p75-IN interaction retargets lentiviral integration.(c) Less elaborate epigenetic heat map comparing epigenetic features known to associated with transcriptionally active or repressive regions (10kb windows). Statistical significance (asterisks, ranked Wald tests) is shown relative to DMSO population (dashes) **p<0.01; ***p<0.001. Significance is reached when p<0.001, compared to MRC. Included features were limited to those identified in high-throughput studies in HeLaP4 and primary $CD4^+$ T-cells. CX014442 treatment during infection induces a dose dependent shift out of transcriptionally active regions. Lower significance is observed at the highest LEDGIN concentrations due to a lower copy number of integration sites. DMSO, Dimethyl sulfoxide.

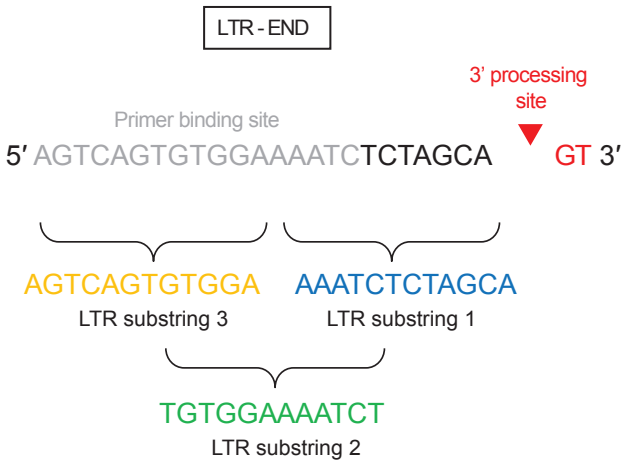


Supplementary Figure S4.4: LEDGIN treatment reduces reactivation from latency.

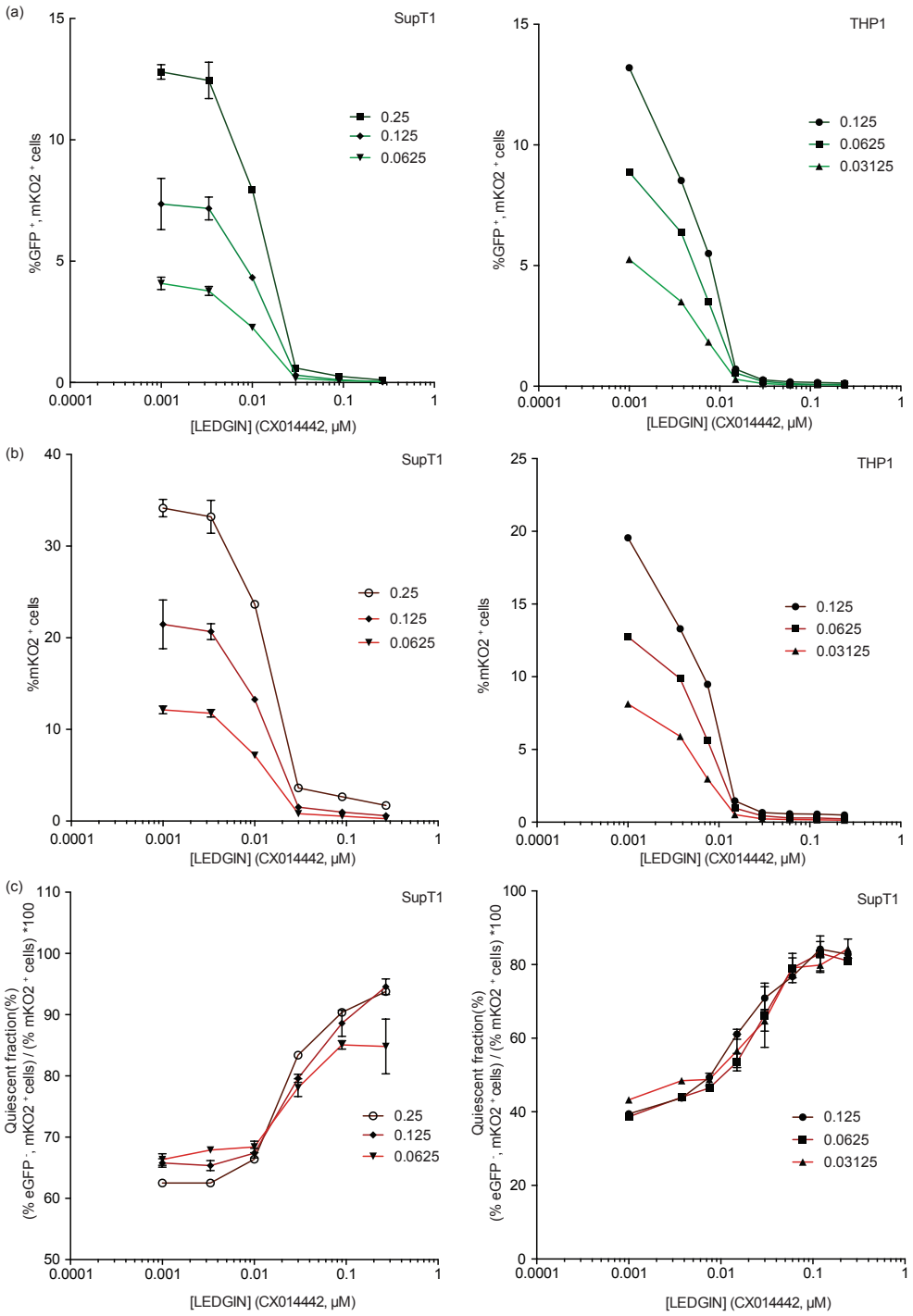
Supplementary Figure S4.4: LEDGIN treatment reduces reactivation from latency. SupT1 cells were infected with single round NL4.3 tCD34 and treated with different concentrations of CX014442 (as indicated). 11 days post infection cells were reactivated using different LRAs at concentrations indicated. (a) % tCD34 positive cells after stimulation with TNFα(10 ng/mL) or DMSO. (b) Fold increase in % tCD34 cells relative to DMSO treatment. All data represent averages of 3 replicates and error bars indicate the standard error of the mean. (c) Bar diagram depicting the fold reactivation (fold increase in % tCD34 positive cells). All data represent averages of 3 replicates and display a representative image for 4 independent experiments. Error bars indicate the standard error of the mean (SEM). A statistical analysis was performed using multiple t-tests and corrected using Sidak-Bonferroni (* $p < 0.05$, ** $p < 0.005$, ***$p < 0.0005$ compared to the DMSO control treatment during infection). (d) Relative fold increase in % tCD34 positive cells (after TNFα stimulation) normalized to untreated control for samples treated with either CX014442 or Raltegravir. (e) Decrease in fraction of quiescent OGH reporter provirus (% $eGFP^-$, $mKO2^+$ cells / % $mKO2^+$ cells *100) upon TNFα stimulation normalized to untreated control for cells treated with either CX014442 or Raltegravir during infection. Actual drug concentrations used were based on either the 'early effect' IC_{50} of CX014442 equal to 3.83 µM or an IC_{50} of Raltegravir equal to 0.0035 µM. SAHA; SuberoylAnilide Hydroxamic Acid, TNFα; Tumor Necrosis Factor alpha, PMA; Phorbol 12- Myristate 13-Acetate, DMSO; DiMethyl SulfOxide, tCD34; truncated Cluster of Differentiation 34. All viruses are VSV-G pseudotyped.



Supplementary Figure S4.5: The integration site sequence logo remains unaffected by LEDGIN-mediated inhibition of the LEDGF/p75-IN interaction. Sequence conservation and relative base frequency in the 18bp genomic DNA sequence surrounding the integration sites (roughly corresponding to the intasome footprint) were evaluated using sequence logos. Sequence conservation is indicated as the total height of each stack (measured in bits), while the relative height of bases in a stack reflects the base frequencies at that position. The site of strand transfer on the plus target DNA strand is depicted as position 0. (a-h) Integration site sequence logos retrieved at different concentrations of LEDGIN CX014442.



Supplementary Figure S4.6: LTR substrings. This example illustrates the process of generating the LTR and LC substrings for substring matching. A lentiviral vector LTR is shown.



Supplementary Figure S4.7: Addition of LEDGINs during production results in transcriptionally silent provirus after integration. HIV-1 double reporter virus (OGH) was produced in HEK293T cells in the presence of varying LEDGIN concentrations. Infection of either SupT1 cells (Left) or THP1 cells (Right) with HIV-1 double reporter virus produced in the presence of increasing concentrations of LEDGIN CX014442 ('late' $IC_{50} = 0.06$ µM). (a) Dose-response curve showing a decrease in % eGFP$^+$, mKO2$^+$ cells with increasing LEDGIN concentration. Three different virus concentrations are depicted in green. (b) Dose-response curve showing a decrease in the overall % mKO2$^+$ cells with increasing LEDGIN concentration. Three different virus concentrations are depicted in red. (c) The fraction of quiescent cells (% eGFP$^-$, mKO2$^+$ cells) / (% mKO2$^+$ cells) *100 increased proportionally with the concentration of LEDGIN added during production. Three different virus concentrations used are shown in red. All viruses are VSV-G pseudotyped. eGFP, Enhanced Green Fluorescent Protein; mKO2, Mutant Kusubira Orange 2.

DataSet	# Sites	% in RefSeq gene	
DMSO	1259	70.37	
CX05045 1x IC_{50}	2693	69.18	
CX05045 5x IC_{50}	791	61.95	***

Supplementary Table S4.1: LEDGIN inhibition of the LEDGF/p75-IN interaction retargets lentiviral integration. Table showing the percentage of HIV integration sites relative to features specific for integration such as integration into the body of genes (Refseq genes). Integration site data sets are obtained from MT4 cells infected with WT HIV (NL4.3) and treated with different concentrations of CX05045. Asterisks depict a significant deviation from the DMSO treated control data set (two-tailed Chi-square test; ***, p-values <0.001).

	DataSet	% Imperfect LTR-Chromosome junctions
LEDGIN (CX014442, µM)	DMSO	1.25
	0.78125	1.32
	1.5625	1.51
	3.125	1.68
	6.25	1.37
	12.5	0.65
	25	1.86
	50	2.58

Supplementary Table S4.2: Percentage of imperfect LTR - chromosome junctions. Non-genomic sequences were detected and removed from the sequence reads to generate the queries for alignment using a Perl substring matching strategy (VISA, Vector Integration Site Analysis server [Hocum et al., 2015]. The percentage imperfect LTR-Chromosome junction was calculated relative to the total number of integration sites for all LEDGIN integration site data sets.

	DataSet	% Integrations in Refseq genes having an opposite orientation	
LEDGIN (CX014442, µM)	DMSO	46.21	-
	0.78125	47.52	*
	1.5625	49.34	*
	3.125	49.83	*
	6.25	49.43	*
	12.5	51.76	**
	25	49.71	*
	50	56.70	***

Supplementary Table S4.3: Integration orientation Table depicting the percentage of integration sites falling within genes that have a similar or opposite orientation relative to the targeted genes. LEDGIN treatment induces a dose-dependent increase in the fraction of integrations having an opposite orientation relative to the target gene. Asterisks depict a significant deviation from the DMSO treated control dataset (Pearsons Chi-square test; ***, p-values <0.005; **, p-values <0.05; *, p-values <0.5).

5 Concluding Discussion

The work presented in this doctoral dissertation encompasses two research manuscripts with the LEDGF/p75-IN interaction forming the common thread throughout the chapters. This discussion frames the different chapters and puts them into perspective.

I) Engineering Safer Next Generation Viral Vectors for Gene Therapeutic Applications.

The capacity to integrate transgenes into the host cell genome makes retroviral vectors interesting tools for gene therapeutic applications where life-long correction is required. Although stable insertion resulted in successful correction of several monogenic disorders, it also accounts for insertional mutagenesis and implies the risk of vector-induced genomic perturbation, major setbacks in otherwise successful clinical gene therapy trials resulting in leukemia development in a subset of treated patients. Despite incremental improvements in vector design, their use is still not risk-free as retroviral vector genotoxicity is largely influenced by the vector integration pattern. A recent retrospective study, however, highlighted the potential of a gene therapy treatment alternative for SCID-X1 as it outperformed an allogenic haploidentical Hematopoietic Stem Cell Transplantation (HSCT) accounting for a faster immune reconstitution and improved thymus response [Touzot et al., 2015]. More importantly, the gene therapy field recently achieved a huge leap forward as the first retroviral vector (Strimvelis^{TM}) was approved in 2016 by the European Commission for the treatment of ADA-SCID [Cicalese et al., 2016]. In order for gene therapy to become a widely applicable first in-line treatment alternative a better understanding of the complex interplay of the parameters responsible for severe adverse events (SAEs) and improved vector genotoxicity profiles are required. At present, there is an increasing interest in new vector platforms displaying a more neutral, close-to-random insertional profile potentially reducing the probability of proto-oncogene activation and lowering the genotoxic potential.
Lentiviral vector (LV) integration is directed into active transcription units by LEDGF/p75, a host-cell protein co-opted by the viral integrase. In this work (see Chapter 3 on page 41), we contributed to the development of safer retroviral vectors and engineered LEDGF/p75-based hybrid tethers in an effort to obtain a more random integration pattern to increase biosafety, and potentially reduce of proto-oncogene activation. We therefore truncated LEDGF/p75 by deleting the N-terminal chromatin-reading PWWP-domain. In parallel we replaced this domain with alternative chromatin binding peptides described in literature as pan-chromatin recognition peptides since they bind cellular chromatin without sequence specificity aiding episomal viruses to persist during mitosis; The spumavirus, Prototype Foamy Virus (PFV), contains a 13-amino acid motif in the group-specific antigen (Gag) binding the H2A/H2B core nucleosome [Tobaly-Tapiero et al., 2008; Nowrouzi et al., 2006; Trobridge et al., 2006]. Likewise, the Kaposi Sarcoma-associated Herpes Virus (KSHV) genome is tethered to the nucleosomal core via a chromatin binding sequence (CBS) at the N-terminal end of the Latency-Associated Nuclear Antigen protein (LANA) [Barbera et al., 2006]. Finally, in the β-Papillomaviruses (PV) a conserved motif in the E2 hinge promotes binding to chromatin and mitotic chromosomes of the invaded cell [Sekhar et al., 2010; Sekhar and McBride, 2012; Võsa et al., 2012]. Each of these hybrid constructs showed unique subnuclear distributions when complemented in LEDGF/p75 depleted cell lines (see Subsection 3.3.2 on

page 46) and expression of these LEDGF-hybrids in LEDGF-depleted cells efficiently rescued LV transduction and integration (see Subsection 3.3.3 on page 48). We characterized proviral integration sites using a nested linker mediated PCR approach and 454 pyrosequencing (see Subsection 3.3.4 on page 50). Analysis revealed LV integrations that distributed more randomly throughout the host-cell genome for the PWWP truncated LEDGF/p75, suggesting the PWWP domain is not required for integration support. The latter phenotype could be attributed to presence of the AT-hook motifs and charged regions present in the N-terminal end of δN_{93}-LEDGF (Cfr. 'scan-and-lock mechanism', [Hendrix et al., 2011]). For most of the peptide constructs (except for PFV $Gag_{534\text{-}546}$-δN_{93}-LEDGF) integration resulted in a more random distribution than under LEDGF/p75 depleted conditions. Only small differences were observed compared to the δN_{93}-LEDGF. The extent of integration targeting however will be dependent on competition with endogenous LEDGF/p75 and will be dependent on the exact levels of endogenous LEDGF/p75 present in the respective target cell. When considering safe harbor criteria, LV integration sites for these LEDGF-hybrids distributed more safely (4-fold increase in % 'safe' sites) compared to LEDGF/p75-mediated integration in wild-type cells. This approach should be broadly applicable to introduce therapeutic or suicide genes for cell therapy, such as patient-specific iPS cells.

Future Perspectives.

Our findings open new options for next generation viral vectors to display a safer genotoxicity profile. In a next step, our approach should be combined with mRNA electroporation of wild-type cells expressing endogenous LEDGF/p75. This method has been previously shown to be very effective in a series of cell types. Alternatively, Hare and colleagues reported on a set of amino-acid substitutions in HIV IN that abolish LEDGF/p75 binding but do not compromise IN activity *in vitro*, together with mutations in the LEDGF/p75 protein that restore binding [Hare et al., 2009]. Engineered viral vectors harboring an artificial IN-LEDGF/p75 pair could facilitate retargeting in a context where WT LEDGF is present. Randomized integration targeting should be directly compared with alpharetroviral vectors displaying an intrinsic, more neutral integration preferences. Ideally, integration is directed towards a single "ultimate" integration environment able to accommodate integration of new genetic material in a manner that enables functionality prediction and does not harm the host cell or organism which is referred to as a "Genomic Safe Harbor" region (GSH) [Papapetrou et al., 2011; Sadelain et al., 2011a; Papapetrou and Schambach, 2016]). However, such regions are yet to be defined. Incremental progress has been made in the tools available to target or edit such unique loci; artificial zinc finger nucleases, meganucleases and TALENs. Fusion of LEDGF to any of these modules has the potential to retarget integration towards these loci. The extent by which integration can be retargeted to a single locus will be limited as integration site selection is additionally affected by more up-stream events such as the nuclear entry route. Therefore it is most likely that targeted integration will display an enrichment of integration events near those specific loci but will also result in multiple off-target integration events. As the gene therapy field moves forward, viral vectors will accumulate additive or synergistic safety features (SIN promoter, insulator sequences, ...) and be evaluated in more general (IVIM [Modlich et al.,

2006, 2009b] or Cdkn2a $^{-/-}$ mouse model [Montini et al., 2006, 2009b]) and disease specific pre-clinical models. The gene therapy field however still lacks powerful assays able to robustly and accurately predict the exact clinical outcome and adverse events. Moreover, standardized parameters to compare different vector platforms should be defined in the future. In addition, it remains unclear whether treatment of polygenic disorders requires different safety criteria. The gene ontology analysis and safe harbor analysis used in this manuscript give us a first indication on the cytotoxic profile of new vector platforms and allow us to check whether the reduced integration frequencies in genes for the respective conditions might result in different GO classes.

II) The Quest for An HIV Cure.

The HIV pandemic represents one of the most important global health challenges in modern history. In the past 30 years, tremendous efforts to develop antiretroviral therapy have led to the development and approval of more than 30 drugs acting on different steps of the HIV replication cycle. cART has revolutionized the treatment of HIV/AIDS dramatically reducing HIV related mortality and morbidity. Yet, anno 2016, cART administration does not cure the patient due to the existence of a latent reservoir of replication competent virus. The presence of this latent reservoir allows HIV to rapidly rebound viremia within weeks after therapy cessation. Moreover, drug related co-morbidities, the economic and therapeutic burden reduce therapy compliance and allow a window of opportunity for HIV to accumulate escape mutations and develop drug resistance. Evidence suggests the development of drug resistance against RAL and EVG (belonging to the potent INSTI class) in treatment-experienced patients [Canducci et al., 2011; Wainberg et al., 2012]. These factors highlight the importance of a continuous effort to identify novel druggable targets and the urge for alternative therapy routes able to exhaust the latent reservoir and control virus replication in the absence of cART, a key priority to the International AIDS Society (IAS) [Deeks et al., 2016].
To date, the latent reservoir remains an insurmountable obstacle to an HIV cure. Early administration of antiretroviral therapy during acute infection limits the size of the latent reservoir but does not completely prevent its establishment. Moreover, cART intensification with the integrase strand transfer inhibitor (INSTI) raltegravir, failed to achieve a further diminution of the reservoir. In pursuit of an HIV cure the main focus aims to purge the latent reservoir in order to elicit a subsequent killing of the reservoir by viral cytopathic effects or immune system clearance while uninfected cells remain protected by cART. An alternative approach proposes to enforce the virus into a 'deep latent' state, an idea that has been lingering in the field but has not been heavily pursued.

LEDGIN-Mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV.

Integration of the viral cDNA into the host cell chromatin catalyzed by IN forms a key step in the viral replication cycle. Targeting of this step led to the development of a potent class of HIV

inhibitors (INSTIs) currently in clinical use. Meanwhile LEDGF/p75 has been thoroughly validated as a key cellular cofactor co-opted by the viral IN for directing HIV integration to the body of actively transcribed genes. The awareness of HIV hijacking endogenous proteins has led to the development of additional inhibitor classes. LEDGINs are *bona fide* small molecules able to abrogate the LEDGF/p75-IN interaction by binding to the IN dimer interface thereby inhibiting HIV replication [Christ et al., 2010]. Surprisingly these compounds perturb the IN dynamics via multiple, interwoven mechanisms. Not only do LEDGINs compete with LEDGF/p75 interaction and thereby block chromatin tethering of the PIC [Christ et al., 2010]; they also inhibit IN catalysis by inducing untimely and unproductive IN assembly [Christ et al., 2012; Kessl et al., 2012; Tsiang et al., 2012; Le Rouzic et al., 2013; Shkriabai et al., 2014; Gupta et al., 2014]. This multimodal mechanism of action results in a steep dose-response curve. Boehringer-Ingelheim (BI) reported on the development of quinoline BI-224436, the first LEDGIN that advanced into clinical trials [Fenwick et al., 2014].

At present, we do not fully understand why, how and when lentiviruses co-opted the chromatin reader LEDGF/p75 during evolution. HIV-1 replication however is severely hampered upon its depletion. Preferential integration of HIV in active transcription units has important implications for the transcriptional activity of the viral promoter and the establishment/maintenance of HIV latency as the local chromatin context is generally believed to affect the transcriptional activity. Here I provide experimental evidence for a strategy to push the virus towards transcriptional quiescence by interference with the LEDGF/p75-IN interaction. In Chapter 4 on page 69 I showed that LEDGF/p75 depletion, known to result in retargeting of integration away from the body of actively transcribed genes [Ciuffi et al., 2005; Shun et al., 2007; Marshall et al., 2007; Schrijvers et al., 2012b] resulted in provirus that was more likely to adopt a transcriptionally quiescent state (Subsection 4.4.1 on page 72). In addition, we demonstrated for the first time that these integrants are less susceptible to reactivation with a panel of latency reversing agents (Subsection 4.4.2 on page 72). Results that provide crucial information for a better understanding of why HIV evolved to use LEDGF/p75 as a cellular cofactor. In addition, I evidenced that LEDGINs (CX014442), acting as potent inhibitors of HIV replication, specifically alter the HIV-1 genomic integration site distribution of the residual integrants in a manner reminiscent of LEDGF/p75 depletion (Subsection 4.4.3 on page 75), a phenotype that was previously unclear. Sharma *et al.* and Jurado *et al.* observed a statistically relevant shift out of gene rich regions upon the addition of the LEDGIN GS-B during infection, a phenotype that could not be reproduced by Gupta or Sharma *et al.* using LEDGIN GSK1264 or KF116 [Sharma et al., 2014; Gupta et al., 2014]. The low number of integration sites or the lower concentrations used (with respect to the IC_{50} of the so-called 'early effect') can possibly explain why statistical significance was not reached in those papers. Co-inhibition of the HRP-2-IN interaction possibly explains the more pronounced shift observed at high CX014442 contentrations. Unexpectedly, a slightly higher fraction of the residual integrants under LEDGIN treatment was integrated in the opposite orientation with respect to the host gene targeted. This phenotype remains a mere observation which possibly contributes to the quiescent phenotype. The underlying mechanism may be related to the induction of aberrant integrase oligomerization which could affect the exact manner by which the intasome interacts with bend/coiled chromatin. However, no effect on the

palindromic sequence logo was observed at increasing LEDGIN concentrations. In addition, LEDGINs also inhibit the integrase-HRP2 interaction possibly responsible for the observed phenotype.
The HIV PIC and provirus are known to locate in proximity to the nuclear rim [Albanese et al., 2008; Di Primio et al., 2013; Marini et al., 2015]. I demonstrated that following LEDGF/p75 depletion and under LEDGIN treatment integration events were located more towards the inner nucleus, away from the nuclear rim or periphery (Subsection 4.4.4 on page 79). Most likely the virus is able to penetrate the chromatin environment for a longer time period before integrating into the host cell chromatin. As previously shown by us and others LEDGINs enhance IN multimerization, inducing a untimely and unproductive IN assembly. My data suggest that addition of CX014442 during the early infection steps and binding to the IN dimer interface does not affect its intrinsic local palindromic sequence preference therefore residual integrations likely represent authentic integration events catalyzed by a WT intasome. Most notably, CX014442 treatment reduced the integration frequency near genomic regions classified as being unsafe based on safe harbor criteria, consistent with the notion that HIV integration does not result in cellular transformation even though some reports linked integration to cellular clonal expansion without oncogenesis. Using a next-generation double-reporter virus (HIV OGH) I evaluated whether a LEDGIN induced shift in integration site distribution could possibly affect the establishment and reactivation of latent provirus. Treatment with LEDGIN, in the first place developed as antivirals, resulted in an overall decrease in the number of infected cells as represented by a decrease in integrated copy number and number of positive cells on FACS analysis. However, the relative fraction of transcriptionally silent provirus compared to the total number of infected cells increased upon increasing concentrations of CX014442 implying that the residual integrants are more likely to adopt a transcriptionally silent state. In line with the phenotype observed for LEDGF/p75 depletion residual integrations under LEDGIN treatment where less susceptible to reactivation by a panel of LRAs in cell lines or general T-cell activating agents in primary $CD4^+$ T-cells. The more pronounced effect observed under LEDGIN treatment could possibily be attributed to the additional inhibition of HRP-2 [Schrijvers et al., 2012a].

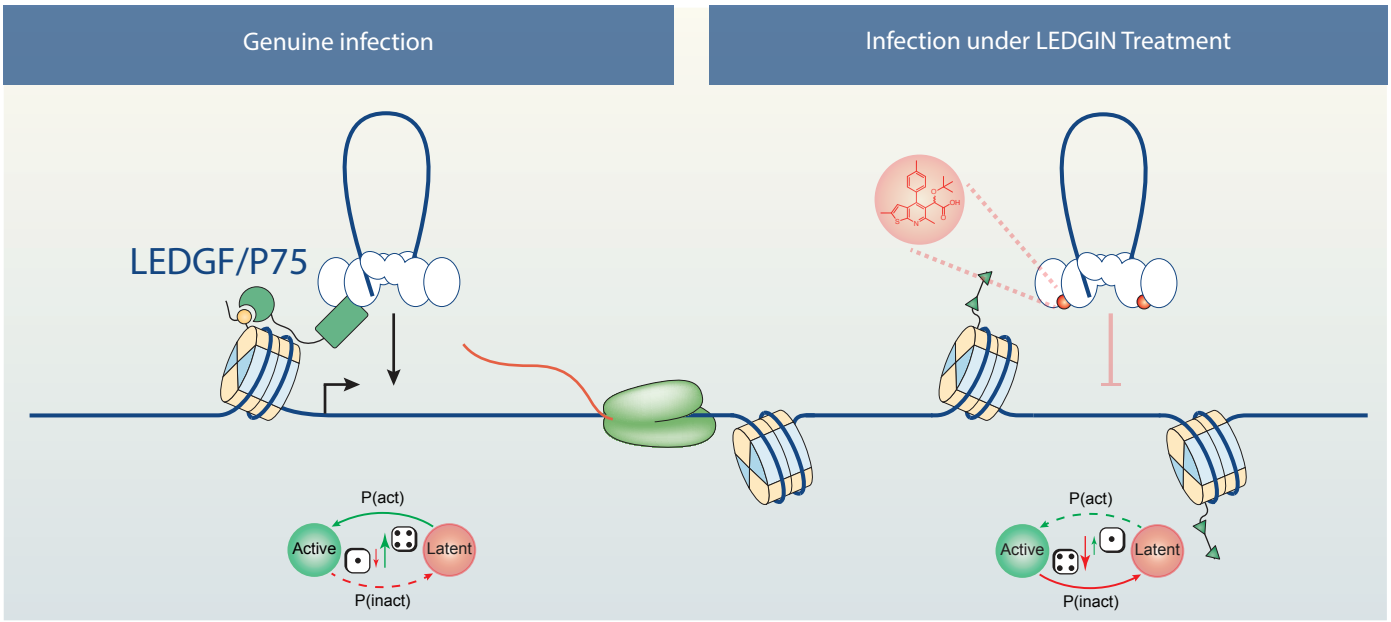


Figure 5.1: Schematic diagram of the effect of LEDGINs on the residual HIV-1 reservoir. LEDGINs, developed as *bona fide* antivirals, result in residual integrants that are more likely to adopt a transcriptionally inactive or quiescent state, less susceptible to reactivation.

Will LEDGIN Molecules Be Able To Play a Role in a Cure for HIV Infection?

HIV reservoirs are believed to be established already at the early Fiebig stages I/II , defined by positive HIV RNA, negative HIV western blot and negative HIV ELISA for stage I plus positive antigenaemia P24 for stage II [McMichael et al., 2009], as early as 10 days after symptoms onset [Ananworanich et al., 2012]. cART provided early during acute infection, drastically reduces the size of the latent reservoir yet is not able to completely prevent reservoir formation, meaning the latent reservoir is, at least in part, is still formed during cART administration. Early treatment initiation with cART therefore will become standard clinical practice in HIV care (INSIGHT START Study Group, 2015) when possible and will notably enhance the recovery of $CD4^+$ T-cell numbers and functions [Hocqueloux et al., 2013]. In addition, early cART impacts the rebound of HIV RNA levels during treatment interruption, allowing greater time off therapy [Wyl et al., 2011]. An approach representing a marked departure from the "shock and kill" paradigm is to enforce viral latency. It was recently proposed that pushing enough proviruses into quiescence could drive the basic reproduction number of HIV below 1, resulting in unsustainable infection [Rouzine et al., 2015]. Suppressing virus production and permanently silence HIV into a so-called "deep latent" state might therefore allow discontinuation of cART administration. Several cases of sustained HIV remission or delayed viral rebound have been previously observed (see Figure 5.2 on page 123) in the setting of HSCT (Berlin and Boston patients, [Henrich et al., 2014]) or early ART administration (Mississippi baby, [Persaud et al., 2013]). In some cases, long-term control of HIV viral loads was observed post-treatment (VISCONTI study in France [Sáez-Cirión et al., 2013], see Figure 5.2). A cART regimen, optimized to significantly and rapidly reduce the latent reservoir size when added early during acute infection and enforcement of a latent phenotype of the residual reser-

voir could potentially hold promise for achieving a long-term HIV remission and lead to less frequent dosing regimens or sustained post-treatment control. Rendering an increased fraction of the ongoing replicating virus inactive upon integration will reduce the possibility of resistance development and postpone the time to relapse. Deep latency could allow the immune system to recover and permanently control the HIV infection as it is promoted to adopt an endogenous retrovirus (ERV)-like state. A combination of LEDGINs together with transcriptional modulators abrogating the recruitment of P-TEFb by the viral Tat protein and impeding HIV production might constitute an attractive treatment option (e.g. Didehydro-Cortistatin A or dCA, [Mousseau et al., 2015]. Tat acts by binding to the TAR element of all viral RNA transcripts and recruits p-TEFb. Inhibition of this interaction will therefore hamper transcription elongation which could render integrated provirus inactive. In comparison to a waddington-epigenetic landscape residual integrants under LEDGIN treatment might display a reduced stochastic transcriptional noise [Dar et al., 2014]. In other words, the threshold levels of Tat required for reactivation might turn out to be higher in comparison to those for genuine integrants. Therefore, residual integrants under a LEDGIN regimen could be more suitable to be locked in a deep latent state by hampering the Tat positive feedback loop. It is clear that this final outcome utterly depends on well-designed clinical trials and surpasses any claims made in this work. Moreover, the outcome of such clinical trials will provide vital information on the time and dynamics of the latent reservoir formation. Well-controlled treatment interruption will contribute to our understanding of the stability of the latent reservoir constituents, and the fraction of individual proviruses responsible for viral re-bound. For HIV pre-exposure prophylaxis (PrEP) LEDGINs have the added benefit that upon PrEP failure or inappropriate PrEP adherence any residual virus that would escape the PrEP treatment is more likely to be detargeted and thus may turn out to be non-productive and refractory to reactivation.

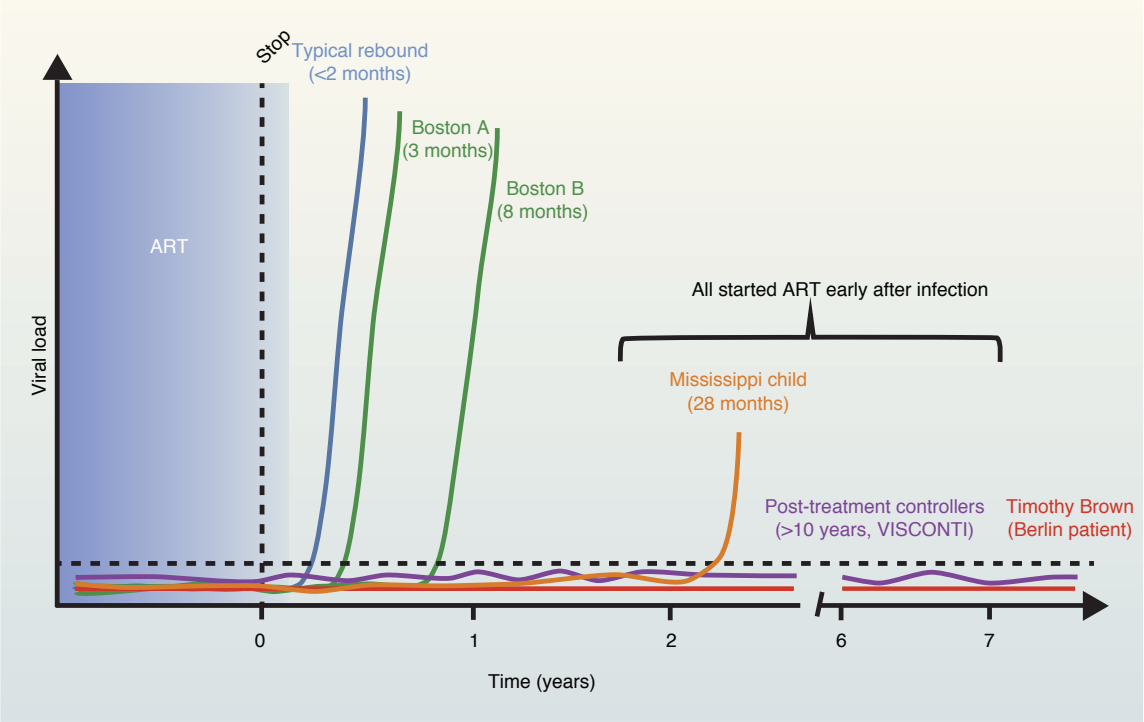


Figure 5.2: Cases of transient or sustained HIV remission off ART. A rapid viral rebound emerges within weeks after therapy cessation. In several cases, viral rebound has been substantially delayed. (*Figure derived from [Deeks et al., 2016]*)

Future perspectives.

To better understand the underlying mechanism by which LEDGIN treatment results in transcriptional silencing of the residual integrants and inhibit reactivation from latency it will be of interest to compare the epigenetic landscape and chromatin environment surrounding the viral LTR promoter after infection under LEDGIN treatment. One could look at the occupancy of Nuc-0 and Nuc-1 histones in the LTR promoter, study CpG methylation profiles or compare histone acetylation/methylation marks using chromatin immune precipitation (ChIP)-based methods and compare these before and after reactivation. For example, acetylation marks and monomethylation marks such as H3K27/K9/K79, H4K20, H2BK5, etc. are known to correlate with active transcription while transcriptionally repressed regions are commonly associated with H3K27me$_3$, H3K9me$_3$ and heterochromatin regions with H4K20me$_3$ and H3K79me$_3$. In addition, a reduced recruitment of RNAP II or host factors supporting HIV transcription, such as NFκB, pTEFβ,... to the LTR promoter might explain the more quiescent phenotype. In parallel, LEDGIN treatment should be compared with alternative antivirals such as raltegravir, elvitegravir and dolutegravir (INSTIs), AZT (RT inhibitor) or ritonavir (protease inhibitor). Treatment with suboptimal doses of the INSTI raltegravir has been reported to lead to aberrant HIV-1 integrations [Varadarajan et al., 2013]. At present, little is known about the exact consequence of aberrant integration events on the transcriptional activity. Current integration site sequencing methods do not allow a large scale study of large LTR deletions/insertions or possible translocations. The SMRTbell circle-sequencing approach of Pacific Biosciences might qualify as an attractive sequencing platform

to study complete LTR signatures as it also provides the possibility to simultaneously detect DNA methylation marks. Preliminary results suggest that LEDGINs added during virus production affect the integration site choice in the next round of infection. The approach employed in this manuscript can be applied in a similar manner to study the effect of LEDGINs added during production on the establishment and maintenance of HIV latency. It will be of paramount importance to extend the observations made in this manuscript to more advanced and clinically relevant primary cell culture models and evaluate the effect of LEDGINs on the size and reactivation potential of the latent reservoir in reservoir constituents other than resting memory $CD4^+$ T-cells. Usage of a WT HIV reporter viruses expressing a fluorescent protein will enable single cell analysis and discriminate between reactivation (increase in % positive cells) and transcriptional activity (MFI differences) in contrast to extracellular p24 measurements (bulk analysis). Studies in animal models such as a humanized mouse model (e.g. the HIS mouse model developed by the French/Swiss company TransCure) or non-human primates [Whitney et al., 2014] will allow follow up the latent reservoir dynamics and provide the possibility to measure the time to-relapse and/or viral set-point upon therapy withdrawal after standard cART, cART + LEDGIN or LEDGIN monotherapy regimens. In the HIS (Human Immune System) mouse model, highly immunodeficient mice are engrafted with purified human cord blood-derived $CD34^+$ cells. 15 weeks after engraftment, mice develop a fully functional human immune system (differentiation of hCD45, and more particularly T-Lymphocytes including $CD4^+$ and $CD8^+$ cells). These mice can be infected with HIV. Recently a mouse viral outgrowth assay was reported showing potential to serve as a powerful tool to identify residual HIV [Metcalf Pate et al., 2015]. It will be of interest to assess multiple anatomical compartments and study the tempo-spatial distribution of replication competent provirus in multiple cell populations. To better understand the exact origin of viral rebound, it will be necessary to implement viral barcoding strategies in proviral integration site sequencing methods (similar to the approach taken in [Akhtar et al., 2013]) combined with a sampling of multiple (not to say all) possible anatomical residencies of latent provirus. Homeostatic proliferation contributing to the latent reservoir might occur less frequent under a LEDGIN regimen as retargeted residual integrants are expected to be less likely to deregulate endogenous gene expression profiles. Compound penetration levels in the lymphoid tissues should be monitored. Final proof, however, will only be obtained in well-designed clinical trials (similar to [Chéret et al., 2015a]). Comparative studies set up to determine the exact levels and ratios of LEDGF/p75 and HRP-2 or other factors contributing to the overall integration preference will provide novel insights in the mechanism responsible for the establishment and maintenance of HIV latency in distinct cell types.

Bibliography

Aiuti, A., Cassani, B., Andolfi, G., Mirolo, M., Biasco, L., Recchia, A., Urbinati, F., Valacca, C., Scaramuzza, S., Aker, M., Slavin, S., Cazzola, M., Sartori, D., Ambrosi, A., Di Serio, C., Roncarolo, M. G., Mavilio, F., and Bordignon, C. (2007). Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. *Journal of Clinical Investigation*, 117(8):2233–2240.

Akhtar, W., de Jong, J., Pindyurin, A. V., Pagie, L., Meuleman, W., de Ridder, J., Berns, A., Wessels, L. F. A., van Lohuizen, M., and van Steensel, B. (2013). Chromatin Position Effects Assayed by Thousands of Reporters Integrated in Parallel. *Cell*, 154(4):914–927.

Albanese, A., Arosio, D., Terreni, M., and Cereseto, A. (2008). HIV-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery. *PLoS ONE*, 3(6):e2413.

Ananworanich, J., Dubé, K., and Chomont, N. (2015). How does the timing of antiretroviral therapy initiation in acute infection affect HIV reservoirs? *Current opinion in HIV and AIDS*, 10(1):18–28.

Ananworanich, J., Schuetz, A., Vandergeeten, C., Sereti, I., de Souza, M., Rerknimitr, R., Dewar, R., Marovich, M., van Griensven, F., Sekaly, R., Pinyakorn, S., Phanuphak, N., Trichavaroj, R., Rutvisuttinunt, W., Chomchey, N., Paris, R., Peel, S., Valcour, V., Maldarelli, F., Chomont, N., Michael, N., Phanuphak, P., Kim, J. H., and RV254/SEARCH 010 Study Group (2012). Impact of multi-targeted antiretroviral treatment on gut T cell depletion and HIV reservoir seeding during acute HIV infection. *PLoS ONE*, 7(3):e33948.

Antoniou, M. N., Skipper, K. A., and Anakok, O. (2013). Optimizing retroviral gene expression for effective therapies. *Human Gene Therapy*, 24(4):363–374.

Archin, N. M., Bateson, R., Tripathy, M. K., Crooks, A. M., Yang, K.-H., Dahl, N. P., Kearney, M. F., Anderson, E. M., Coffin, J. M., Strain, M. C., Richman, D. D., Robertson, K. R., Kashuba, A. D., Bosch, R. J., Hazuda, D. J., Kuruc, J. D., Eron, J. J., and Margolis, D. M. (2014a). HIV-1 expression within resting CD4+ T cells after multiple doses of vorinostat. *The Journal of infectious diseases*, 210(5):728–735.

Archin, N. M., Liberty, A. L., Kashuba, A. D., Choudhary, S. K., Kuruc, J. D., Crooks, A. M., Parker, D. C., Anderson, E. M., Kearney, M. F., Strain, M. C., Richman, D. D., Hudgens, M. G., Bosch, R. J., Coffin, J. M., Eron, J. J., Hazuda, D. J., and Margolis, D. M. (2012). Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature*, 487(7408):482–485.

Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N., and Margolis, D. M. (2014b). Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nature Reviews Microbiology*, 12(11):750–764.

Balakrishnan, M., Yant, S. R., Tsai, L., O'Sullivan, C., Bam, R. A., Tsai, A., Niedziela-Majka, A., Stray, K. M., Sakowicz, R., and Cihlar, T. (2013). Non-catalytic site HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in target cells. *PLoS ONE*, 8(9):e74163.

Ballandras-Colas, A., Brown, M., Cook, N. J., Dewdney, T. G., Demeler, B., Cherepanov, P., Lyumkis, D., and Engelman, A. N. (2016). Cryo-EM reveals a novel octameric integrase structure for betaretroviral intasome function. *Nature*, 530(7590):358–361.

Barbera, A. J., Chodaparambil, J. V., Kelley-Clarke, B., Joukov, V., Walter, J. C., Luger, K., and Kaye, K. M. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science*, 311(5762):856–861.

Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, 220(4599):868–871.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, 129(4):823–837.

Bartholomeeusen, K., Christ, F., Hendrix, J., Rain, J.-C., Emiliani, S., Benarous, R., Debyser, Z., Gijsbers, R., and De Rijck, J. (2009). Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PogZ. *Journal of Biological Chemistry*, 284(17):11467–11477.

Bartholomeeusen, K., De Rijck, J., Busschots, K., Desender, L., Gijsbers, R., Emiliani, S., Benarous, R., Debyser, Z., and Christ, F. (2007). Differential Interaction of HIV-1 Integrase and JPO2 with the C Terminus of LEDGF/p75. *Journal of Molecular Biology*, 372(2):407–421.

Bartholomeeusen, K., Fujinaga, K., Xiang, Y., and Peterlin, B. M. (2013). Histone deacetylase inhibitors (HDACis) that release the positive transcription elongation factor b (P-TEFb) from its inhibitory complex also activate HIV transcription. *The Journal of biological chemistry*, 288(20):14400–14407.

Barton, K., Winckelmann, A., and Palmer, S. (2016). HIV-1 Reservoirs During Suppressive Therapy. *Trends in microbiology*, 24(5):345–355.

Barton, K. M., Archin, N. M., Keedy, K. S., Espeseth, A. S., Zhang, Y.-l., Gale, J., Wagner, F. F., Holson, E. B., and Margolis, D. M. (2014). Selective HDAC inhibition for the disruption of latent HIV-1 infection. *PLoS ONE*, 9(8):e102684.

Blaese, R. M., Culver, K. W., Miller, A. D., Carter, C. S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P., Greenblatt, J. J., Rosenberg, S. A., Klein, H., Berger, M., Mullen, C. A., Ramsey, W. J., Muul, L., Morgan, R. A., and Anderson, W. F. (1995). T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science*, 270(5235):475–480.

Blazkova, J., Murray, D., Justement, J. S., Funk, E. K., Nelson, A. K., Moir, S., Chun, T.-W., and Fauci, A. S. (2012). Paucity of HIV DNA methylation in latently infected, resting CD4+ T cells from infected individuals receiving antiretroviral therapy. *Journal of Virology*, 86(9):5390–5392.

Blazkova, J., Trejbalova, K., Gondois-Rey, F., Halfon, P., Philibert, P., Guiguen, A., Verdin, E., Olive, D., Van Lint, C., Hejnar, J., and Hirsch, I. (2009). CpG methylation controls reactivation of HIV from latency. *PLoS Pathogens*, 5(8):e1000554.

Bordignon, C., Notarangelo, L. D., Nobili, N., Ferrari, G., Casorati, G., Panina, P., Mazzolari, E., Maggioni, D., Rossi, C., Servida, P., Ugazio, A. G., and Mavilio, F. (1995). Gene therapy in peripheral blood lymphocytes and bone marrow for ADA- immunodeficient patients. *Science*, 270(5235):470–475.

Borrenberghs, D., Thys, W., Rocha, S., Demeulemeester, J., Weydert, C., Dedecker, P., Hofkens, J., Debyser, Z., and Hendrix, J. (2014). HIV virions as nanoscopic test tubes for probing oligomerization of the integrase enzyme. *ACS nano*, 8(4):3531–3545.

Bosque, A., Famiglietti, M., Weyrich, A. S., Goulston, C., and Planelles, V. (2011). Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells. *PLoS Pathogens*, 7(10):e1002288.

Bosticardo, M., Marangoni, F., Aiuti, A., Villa, A., and Grazia Roncarolo, M. (2009). Recent advances in understanding the pathophysiology of Wiskott-Aldrich syndrome. *Blood*, 113(25):6288–6295.

Bouchat, S., Gatot, J.-S., Kabeya, K., Cardona, C., Colin, L., Herbein, G., De Wit, S., Clumeck, N., Lambotte, O., Rouzioux, C., Rohr, O., and Van Lint, C. (2012). Histone methyltransferase inhibitors induce HIV-1 recovery in resting CD4+ T cells from HIV-1-infected HAART-treated patients. *AIDS*, 26(12):1473–1482.

Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P. P., Díez, I. A., Dewey, R. A., Böhm, M., Nowrouzi, A., Ball, C. R., Glimm, H., Naundorf, S., Kühlcke, K., Blasczyk, R., Kondratenko, I., Maródi, L., Orange, J. S., von Kalle, C., and Klein, C. (2010). Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *The New England journal of medicine*, 363(20):1918–1927.

Braun, C. J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Göhring, G., Steinemann, D., Fronza, R., Ball, C. R., Haemmerle, R., Naundorf, S., Kühlcke, K., Rose, M., Fraser, C., Mathias, L., Ferrari, R., Abboud, M. R., Al-Herz, W., Kondratenko, I., Maródi, L., Glimm, H., Schlegelberger, B., Schambach, A., Albert, M. H., Schmidt, M., von Kalle, C., and Klein, C. (2014). Gene therapy for Wiskott-Aldrich syndrome–long-term efficacy and genotoxicity. *Science translational medicine*, 6(227):227ra33–227ra33.

Bruner, K. M., Hosmane, N. N., and Siliciano, R. F. (2015). Towards an HIV-1 cure: measuring the latent reservoir. *Trends in microbiology*, 23(4):192–203.

Budhiraja, S., Famiglietti, M., Bosque, A., Planelles, V., and Rice, A. P. (2013). Cyclin T1 and CDK9 T-loop phosphorylation are downregulated during establishment of HIV-1 latency in primary resting memory CD4+ T cells. *Journal of Virology*, 87(2):1211–1220.

Burnett, J. C., Miller-Jensen, K., Shah, P. S., Arkin, A. P., and Schaffer, D. V. (2009). Control of stochastic gene expression by host factors at the HIV promoter. *PLoS Pathogens*, 5(1):e1000260.

Bushman, F. D. (1994). Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proceedings of the National Academy of Sciences*, 91(20):9233–9237.

Bushman, F. D. and Miller, M. D. (1997). Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *Journal of Virology*, 71(1):458–464.

Butler, S. L., Hansen, M. S., and Bushman, F. D. (2001). A quantitative assay for HIV DNA integration in vivo. *Nature Medicine*, 7(5):631–634.

Buzón, M. J., Massanella, M., Llibre, J. M., Esteve, A., Dahl, V., Puertas, M. C., Gatell, J. M., Domingo, P., Paredes, R., Sharkey, M., Palmer, S., Stevenson, M., Clotet, B., Blanco, J., and Martinez-Picado, J. (2010). HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. *Nature Medicine*, 16(4):460–465.

Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1997). Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nature structural biology*, 4(7):567–577.

Calvanese, V., Chavez, L., Laurent, T., Ding, S., and Verdin, E. (2013). Dual-color HIV reporters trace a population of latently infected cells and enable their purification. *Virology*, 446(1-2):283–292.

Canducci, F., Ceresola, E. R., Boeri, E., Spagnuolo, V., Cossarini, F., Castagna, A., Lazzarin, A., and Clementi, M. (2011). Cross-resistance profile of the novel integrase inhibitor Dolutegravir (S/GSK1349572) using clonal viral variants selected in patients failing raltegravir. *The Journal of infectious diseases*, 204(11):1811–1815.

Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C. C., Veres, G., Schmidt, M., Kutschera, I., Vidaud, M., Abel, U., Dal Cortivo, L., Caccavelli, L., Mahlaoui, N., Kiermer, V., Mittelstaedt, D., Bellesme, C., Lahlou, N., Lefrère, F., Blanche, S., Audit, M., Payen, E., Leboulch, P., l'Homme, B., Bougnères, P., von Kalle, C., Fischer, A., Cavazzana-Calvo, M., and Aubourg, P. (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science*, 326(5954):818–823.

Cattoglio, C., Facchini, G., Sartori, D., Antonelli, A., Miccio, A., Cassani, B., Schmidt, M., von Kalle, C., Howe, S., Thrasher, A. J., Aiuti, A., Ferrari, G., Recchia, A., and Mavilio, F. (2007). Hot spots of retroviral integration in human CD34+ hematopoietic cells. *Blood*, 110(6):1770–1778.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J. L., Bousso, P., Deist, F. L., and Fischer, A. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science*, 288(5466):669–672.

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., Cavallesco, R., Gillet-Legrand, B., Caccavelli, L., Sgarra, R., Maouche-Chrétien, L., Bernaudin, F., Girot, R., Dorazio, R., Mulder, G.-J., Polack, A., Bank, A., Soulier, J., Larghero, J., Kabbara, N., Dalle, B., Gourmel, B., Socie, G., Chrétien, S., Cartier, N., Aubourg, P., Fischer, A., Cornetta, K., Galacteros, F., Beuzard, Y., Gluckman, E., Bushman, F., Hacein-Bey-Abina, S., and Leboulch, P. (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. *Nature*, 467(7313):318–322.

Centers for Disease Control (CDC) (1981). Pneumocystis pneumonia–Los Angeles. *MMWR. Morbidity and mortality weekly report*, 30(21):250–252.

Cesana, D., Ranzani, M., Volpin, M., Bartholomae, C., Duros, C., Artus, A., Merella, S., Benedicenti, F., Sergi Sergi, L., Sanvito, F., Brombin, C., Nonis, A., Serio, C. D., Doglioni, C., von Kalle, C., Schmidt, M., Cohen-Haguenauer, O., Naldini, L., and Montini, E. (2014). Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. *Molecular therapy : the journal of the American Society of Gene Therapy*, 22(4):774–785.

Cesana, D., Sgualdino, J., Rudilosso, L., Merella, S., Naldini, L., and Montini, E. (2012). Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. *Journal of Clinical Investigation*, 122(5):1667–1676.

Chatziandreou, I., Siapati, E. K., and Vassilopoulos, G. (2011). Genetic correction of X-linked chronic granulomatous disease with novel foamy virus vectors. *Experimental hematology*, 39(6):643–652.

Chavez, L., Calvanese, V., and Verdin, E. (2015). HIV Latency Is Established Directly and Early in Both Resting and Activated Primary CD4 T Cells. *PLoS Pathogens*, 11(6):e1004955.

Cherepanov, P., Devroe, E., Silver, P. A., and Engelman, A. (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *Journal of Biological Chemistry*, 279(47):48883–48892.

Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *Journal of Biological Chemistry*, 278(1):372–381.

Cherepanov, P., Maertens, G. N., and Hare, S. (2011). Structural insights into the retroviral DNA integration apparatus. *Current opinion in structural biology*, 21(2):249–256.

Cherepanov, P., Sun, Z.-Y. J., Rahman, S., Maertens, G., Wagner, G., and Engelman, A. (2005). Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. *Nature Structural & Molecular Biology*, 12(6):526–532.

Chéret, A., Bacchus-Souffan, C., Avettand-Fènoël, V., Melard, A., Nembot, G., Blanc, C., Samri, A., Sáez-Cirión, A., Hocqueloux, L., Lascoux-Combe, C., Allavena, C., Goujard, C., Valantin, M. A., Leplatois, A., Meyer, L., Rouzioux, C., Autran, B., and OPTIPRIM ANRS-147 Study Group (2015a). Combined ART started during acute HIV infection protects central memory CD4+ T cells and can induce remission. *The Journal of antimicrobial chemotherapy*, 70(7):2108–2120.

Chéret, A., Nembot, G., Melard, A., Lascoux, C., Slama, L., Miailhes, P., Yeni, P., Abel, S., Avettand-Fènoël, V., Venet, A., Chaix, M.-L., Molina, J.-M., Katlama, C., Goujard, C., Tamalet, C., Raffi, F., Lafeuillade, A., Reynes, J., Ravaux, I., Hoen, B., Delfraissy, J.-F., Meyer, L., Rouzioux, C., and OPTIPRIM ANRS Study Group (2015b). Intensive five-drug antiretroviral therapy regimen versus standard triple-drug therapy during primary HIV-1 infection (OPTIPRIM-ANRS 147): a randomised, open-label, phase 3 trial. *The Lancet. Infectious diseases*, 15(4):387–396.

Chiang, K. and Rice, A. P. (2012). MicroRNA-mediated restriction of HIV-1 in resting CD4+ T cells and monocytes. *Viruses*, 4(9):1390–1409.

Chomont, N., El-Far, M., Ancuta, P., Trautmann, L., Procopio, F. A., Yassine-Diab, B., Boucher, G., Boulassel, M.-R., Ghattas, G., Brenchley, J. M., Schacker, T. W., Hill, B. J., Douek, D. C., Routy, J.-P., Haddad, E. K., and Sékaly, R.-P. (2009). HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature Medicine*, 15(8):893–900.

Christ, F. and Debyser, Z. (2013). The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology*, 435(1):102–109.

Christ, F., Shaw, S., Demeulemeester, J., Desimmie, B. A., Marchand, A., Butler, S., Smets, W., Chaltin, P., Westby, M., Debyser, Z., and Pickford, C. (2012). Small-molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. *Antimicrobial Agents and Chemotherapy*, 56(8):4365–4374.

Christ, F., Voet, A., Marchand, A., Nicolet, S., Desimmie, B. A., Marchand, D., Bardiot, D., Van der Veken, N. J., Van Remoortel, B., Strelkov, S. V., De Maeyer, M., Chaltin, P., and Debyser, Z. (2010). Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nature Chemical Biology*, 6(6):442–448.

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., Kuo, Y. H., Brookmeyer, R., Zeiger, M. A., Barditch-Crovo, P., and Siliciano, R. F. (1997a). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, 387(6629):183–188.

Chun, T.-W., Nickle, D. C., Justement, J. S., Meyers, J. H., Roby, G., Hallahan, C. W., Kottilil, S., Moir, S., Mican, J. M., Mullins, J. I., Ward, D. J., Kovacs, J. A., Mannon, P. J., and Fauci, A. S. (2008). Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy. *The Journal of infectious diseases*, 197(5):714–720.

Chun, T. W., Stuyver, L., Mizell, S. B., Ehler, L. A., Mican, J. A., Baseler, M., Lloyd, A. L., Nowak, M. A., and Fauci, A. S. (1997b). Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proceedings of the National Academy of Sciences of the United States of America*, 94(24):13193–13197.

Churchill, M. J., Wesselingh, S. L., Cowley, D., Pardo, C. A., McArthur, J. C., Brew, B. J., and Gorry, P. R. (2009). Extensive astrocyte infection is prominent in human immunodeficiency virus-associated dementia. *Annals of neurology*, 66(2):253–258.

Cicalese, M. P., Ferrua, F., Castagnaro, L., Pajno, R., Barzaghi, F., Giannelli, S., Dionisio, F., Brigida, I., Bonopane, M., Casiraghi, M., Tabucchi, A., Carlucci, F., Grunebaum, E., Adeli, M., Bredius, R. G., Puck, J. M., Stepensky, P., Tezcan, I., Rolfe, K., De Boever, E., Reinhardt, R. R., Appleby, J., Ciceri, F., Roncarolo, M. G., and Aiuti, A. (2016). Update on the safety and efficacy of retroviral gene therapy for immunodeficiency due to adenosine deaminase deficiency. *Blood*.

Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. *Nature Medicine*, 11(12):1287–1289.

Dahabieh, M. S., Battivelli, E., and Verdin, E. (2015). Understanding HIV latency: the road to an HIV cure. *Annual review of medicine*, 66(1):407–421.

Dahabieh, M. S., Ooms, M., Simon, V., and Sadowski, I. (2013). A doubly fluorescent HIV-1 reporter shows that the majority of integrated HIV-1 is latent shortly after infection. *Journal of Virology*, 87(8):4716–4727.

Dar, R. D., Hosmane, N. N., Arkin, M. R., Siliciano, R. F., and Weinberger, L. S. (2014). Screening for noise in gene expression identifies drug synergies. *Science*, 344(6190):1392–1396.

Dar, R. D., Razooky, B. S., Singh, A., Trimeloni, T. V., McCollum, J. M., Cox, C. D., Simpson, M. L., and Weinberger, L. S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, 109(43):17454–17459.

Darcis, G., Kula, A., Bouchat, S., Fujinaga, K., Corazza, F., Ait-Ammar, A., Delacourt, N., Melard, A., Kabeya, K., Vanhulle, C., Van Driessche, B., Gatot, J.-S., Cherrier, T., Pianowski, L. F., Gama, L., Schwartz, C., Vila, J., Burny, A., Clumeck, N., Moutschen, M., De Wit, S., Peterlin, B. M., Rouzioux, C., Rohr, O., and Van Lint, C. (2015). An In-Depth Comparison of Latency-Reversing Agent Combinations in Various In Vitro and Ex Vivo HIV-1 Latency Models Identified Bryostatin-1+JQ1 and Ingenol-B+JQ1 to Potently Reactivate Viral Gene Expression. *PLoS Pathogens*, 11(7):e1005063.

Daugaard, M., Baude, A., Fugger, K., Povlsen, L. K., Beck, H., Sørensen, C. S., Petersen, N. H. T., Sorensen, P. H. B., Lukas, C., Bartek, J., Lukas, J., Rohde, M., and Jäättelä, M. (2012). LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nature Structural & Molecular Biology*, 19(8):803–810.

De Ravin, S. S., Su, L., Theobald, N., Choi, U., Macpherson, J. L., Poidinger, M., Symonds, G., Pond, S. M., Ferris, A. L., Hughes, S. H., Malech, H. L., and Wu, X. (2014). Enhancers are major targets for murine leukemia virus vector integration. *Journal of Virology*, 88(8):4504–4513.

De Rijck, J., Bartholomeeusen, K., Ceulemans, H., Debyser, Z., and Gijsbers, R. (2010). High-resolution profiling of the LEDGF/p75 chromatin interaction in the ENCODE region. *Nucleic Acids Research*, 38(18):6135–6147.

De Rijck, J., de Kogel, C., Demeulemeester, J., Vets, S., El Ashkar, S., Malani, N., Bushman, F. D., Landuyt, B., Husson, S. J., Busschots, K., Gijsbers, R., and Debyser, Z. (2013). The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration near Transcription Start Sites. *CellReports*, 5(4):886–894.

De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F., and Debyser, Z. (2006). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *Journal of Virology*, 80(23):11498–11509.

Debyser, Z., Christ, F., De Rijck, J., and Gijsbers, R. (2015). Host factors for retroviral integration site selection. *Trends in biochemical sciences*, 40(2):108–116.

Deeks, S. G., Lewin, S. R., Ross, A. L., Ananworanich, J., Benkirane, M., Cannon, P., Chomont, N., Douek, D., Lifson, J. D., Lo, Y.-R., Kuritzkes, D., Margolis, D., Mellors, J., Persaud, D., Tucker, J. D., Barré-Sinoussi, F., International AIDS Society Towards a Cure Working Group, Alter, G., Auerbach, J., Autran, B., Barouch, D. H., Behrens, G., Cavazzana, M., Chen, Z., Cohen, É. A., Corbelli, G. M., Eholié, S., Eyal, N., Fidler, S., Garcia, L., Grossman, C., Henderson, G., Henrich, T. J., Jefferys, R., Kiem, H.-P., McCune, J., Moodley, K., Newman, P. A., Nijhuis, M., Nsubuga, M. S., Ott, M., Palmer, S., Richman, D., Sáez-Cirión, A., Sharp, M., Siliciano, J., Silvestri, G., Singh, J., Spire, B., Taylor, J., Tolstrup, M., Valente, S., van Lunzen, J., Walensky, R., Wilson, I., and Zack, J. (2016). International AIDS Society global scientific strategy: towards an HIV cure 2016. *Nature Medicine*.

Deichmann, A., Hacein-Bey-Abina, S., Schmidt, M., Garrigue, A., Brugman, M. H., Hu, J., Glimm, H., Gyapay, G., Prum, B., Fraser, C. C., Fischer, N., Schwarzwaelder, K., Siegler, M.-L., de Ridder, D., Pike-Overzet, K., Howe, S. J., Thrasher, A. J., Wagemaker, G., Abel, U., Staal, F. J. T., Delabesse, E., Villeval, J.-L., Aronow, B., Hue, C., Prinz, C., Wissler, M., Klanke, C., Weissenbach, J., Alexander, I., Fischer, A., von Kalle, C., and Cavazzana-Calvo, M. (2007). Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. *Journal of Clinical Investigation*, 117(8):2225–2232.

Demeulemeester, J., Chaltin, P., Marchand, A., De Maeyer, M., Debyser, Z., and Christ, F. (2014a). LEDGINs, non-catalytic site inhibitors of HIV-1 integrase: a patent review (2006 - 2014). *Expert opinion on therapeutic patents*, 24(6):609–632.

Demeulemeester, J., De Rijck, J., Gijsbers, R., and Debyser, Z. (2015). Retroviral integration: Site matters: Mechanisms and consequences of retroviral integration site selection. *BioEssays*, pages n/a–n/a.

Demeulemeester, J., Vets, S., Schrijvers, R., Madlala, P., De Maeyer, M., De Rijck, J., Ndung'u, T., Debyser, Z., and Gijsbers, R. (2014b). HIV-1 Integrase Variants Retarget Viral Integration and Are Associated with Disease Progression in a Chronic Infection Cohort. *Cell Host and Microbe*, 16(5):651–662.

Deng, K., Pertea, M., Rongvaux, A., Wang, L., Durand, C. M., Ghiaur, G., Lai, J., McHugh, H. L., Hao, H., Zhang, H., Margolick, J. B., Gurer, C., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Deeks, S. G., Strowig, T., Kumar, P., Siliciano, J. D., Salzberg, S. L., Flavell, R. A., Shan, L., and Siliciano, R. F. (2015). Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature*, 517(7534):381–385.

Derse, D., Crise, B., Li, Y., Princler, G., Lum, N., Stewart, C., McGrath, C. F., Hughes, S. H., Munroe, D. J., and Wu, X. (2007). Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *Journal of Virology*, 81(12):6731–6741.

Desimmie, B. A., Schrijvers, R., Demeulemeester, J., Borrenberghs, D., Weydert, C., Thys, W., Vets, S., Van Remoortel, B., Hofkens, J., De Rijck, J., Hendrix, J., Bannert, N., Gijsbers, R., Christ, F., and Debyser, Z. (2013). LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the virions. *Retrovirology*, 10(1):57.

Di Primio, C., Quercioli, V., Allouch, A., Gijsbers, R., Christ, F., Debyser, Z., Arosio, D., and Cereseto, A. (2013). Single-cell imaging of HIV-1 provirus (SCIP). *Proceedings of the National Academy of Sciences of the United States of America*, 110(14):5636–5641.

Eidahl, J. O., Crowe, B. L., North, J. A., McKee, C. J., Shkriabai, N., Feng, L., Plumb, M., Graham, R. L., Gorelick, R. J., Hess, S., Poirier, M. G., Foster, M. P., and Kvaratskhelia, M. (2013). Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Research*, 41(6):3924–3936.

Eijkelenboom, A. P., van den Ent, F. M., Wechselberger, R., Plasterk, R. H., Kaptein, R., and Boelens, R. (2000). Refined solution structure of the dimeric N-terminal HHCC domain of HIV-2 integrase. *Journal of biomolecular NMR*, 18(2):119–128.

El Ashkar, S., De Rijck, J., Demeulemeester, J., Vets, S., Madlala, P., Cermakova, K., Debyser, Z., and Gijsbers, R. (2014). BET-independent MLV-based Vectors Target Away From Promoters and Regulatory Elements. *Molecular Therapy—Nucleic Acids*, 3(7):e179.

Elis, E., Ehrlich, M., Prizan-Ravid, A., Laham-Karam, N., and Bacharach, E. (2012). p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. *PLoS Pathogens*, 8(12):e1003103.

Elliott, J. H., Wightman, F., Solomon, A., Ghneim, K., Ahlers, J., Cameron, M. J., Smith, M. Z., Spelman, T., McMahon, J., Velayudham, P., Brown, G., Roney, J., Watson, J., Prince, M. H., Hoy, J. F., Chomont, N., Fromentin, R., Procopio, F. A., Zeidan, J., Palmer, S., Odevall, L., Johnstone, R. W., Martin, B. P., Sinclair, E., Deeks, S. G., Hazuda, D. J., Cameron, P. U., Sékaly, R.-P., and Lewin, S. R. (2014). Activation of HIV transcription with short-course vorinostat in HIV-infected patients on suppressive antiretroviral therapy. *PLoS Pathogens*, 10(10):e1004473.

Ellis, J. (2005). Silencing and variegation of gammaretrovirus and lentivirus vectors. *Human Gene Therapy*, 16(11):1241–1246.

Emery, D. W. (2011). The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. *Human Gene Therapy*, 22(6):761–774.

Engelman, A. and Cherepanov, P. (2014). Retroviral Integrase Structure and DNA Recombination Mechanism. *Microbiology spectrum*, 2(6):1011–1033.

Eriksson, S., Graf, E. H., Dahl, V., Strain, M. C., Yukl, S. A., Lysenko, E. S., Bosch, R. J., Lai, J., Chioma, S., Emad, F., Abdel-Mohsen, M., Hoh, R., Hecht, F., Hunt, P., Somsouk, M., Wong, J., Johnston, R., Siliciano, R. F., Richman, D. D., O'Doherty, U., Palmer, S., Deeks, S. G., and Siliciano, J. D. (2013). Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathogens*, 9(2):e1003174.

Fadel, H. J., Morrison, J. H., Saenz, D. T., Fuchs, J. R., Kvaratskhelia, M., Ekker, S. C., and Poeschla, E. M. (2014). TALEN knockout of the PSIP1 gene in human cells: analyses of HIV-1 replication and allosteric integrase inhibitor mechanism. *Journal of Virology*, 88(17):9704–9717.

Fader, L. D., Malenfant, E., Parisien, M., Carson, R., Bilodeau, F., Landry, S., Pesant, M., Brochu, C., Morin, S., Chabot, C., Halmos, T., Bousquet, Y., Bailey, M. D., Kawai, S. H., Coulombe, R., LaPlante, S., Jakalian, A., Bhardwaj, P. K., Wernic, D., Schroeder, P., Amad, M., Edwards, P., Garneau, M., Duan, J., Cordingley, M., Bethell, R., Mason, S. W., Bös, M., Bonneau, P., Poupart, M.-A., Faucher, A.-M., Simoneau, B., Fenwick, C., Yoakim, C., and Tsantrizos, Y. (2014). Discovery of BI 224436, a Noncatalytic Site Integrase Inhibitor (NCINI) of HIV-1. *ACS medicinal chemistry letters*, 5(4):422–427.

Fehse, B., Richters, A., Putimtseva-Scharf, K., Klump, H., Li, Z., Ostertag, W., Zander, A. R., and Baum, C. (2000). CD34 splice variant: an attractive marker for selection of gene-modified cells. *Molecular Therapy*, 1(5 Pt 1):448–456.

Fenwick, C., Amad, M., Bailey, M. D., Bethell, R., Bös, M., Bonneau, P., Cordingley, M., Coulombe, R., Duan, J., Edwards, P., Fader, L. D., Faucher, A.-M., Garneau, M., Jakalian, A., Kawai, S., Lamorte, L., LaPlante, S., Luo, L., Mason, S., Poupart, M.-A., Rioux, N., Schroeder, P., Simoneau, B., Tremblay, S., Tsantrizos, Y., Witvrouw, M., and Yoakim, C. (2014). Preclinical profile of BI 224436, a novel HIV-1 non-catalytic-site integrase inhibitor. *Antimicrobial Agents and Chemotherapy*, 58(6):3233–3244.

Ferris, A. L., Wu, X., Hughes, C. M., Stewart, C., Smith, S. J., Milne, T. A., Wang, G. G., Shun, M.-C., Allis, C. D., Engelman, A., and Hughes, S. H. (2010). Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proceedings of the National Academy of Sciences of the United States of America*, 107(7):3135–3140.

Finzi, D., Blankson, J., Siliciano, J. D., Margolick, J. B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T. C., Chaisson, R. E., Rosenberg, E., Walker, B., Gange, S., Gallant, J., and Siliciano, R. F. (1999). Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature Medicine*, 5(5):512–517.

Finzi, D., Hermankova, M., Pierson, T., Carruth, L. M., Buck, C., Chaisson, R. E., Quinn, T. C., Chadwick, K., Margolick, J., Brookmeyer, R., Gallant, J., Markowitz, M., Ho, D. D., Richman, D. D., and Siliciano, R. F. (1997). Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*, 278(5341):1295–1300.

Fletcher, C. V., Staskus, K., Wietgrefe, S. W., Rothenberger, M., Reilly, C., Chipman, J. G., Beilman, G. J., Khoruts, A., Thorkelson, A., Schmidt, T. E., Anderson, J., Perkey, K., Stevenson, M., Perelson, A. S., Douek, D. C., Haase, A. T., and Schacker, T. W. (2014). Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6):2307–2312.

Foley, B. T., Leitner, T. K., Apetrei, C., Hahn, B., Mizrachi, I., Mullins, J., Rambaut, A., Wolinsky, S., and Korber, B. T. M. (2015). HIV sequence compendium 2015. Technical report.

Francis, A. C., Di Primio, C., Quercioli, V., Valentini, P., Boll, A., Girelli, G., Demichelis, F., Arosio, D., and Cereseto, A. (2014). Second generation imaging of nuclear/cytoplasmic HIV-1 complexes. *AIDS research and human retroviruses*, 30(7):717–726.

Frenkel, L. M., Wang, Y., Learn, G. H., McKernan, J. L., Ellis, G. M., Mohan, K. M., Holte, S. E., De Vange, S. M., Pawluk, D. M., Melvin, A. J., Lewis, P. F., Heath, L. M., Beck, I. A., Mahalanabis, M., Naugler, W. E., Tobin, N. H., and Mullins, J. I. (2003). Multiple viral genetic analyses detect low-level human immunodeficiency virus type 1 replication during effective highly active antiretroviral therapy. *Journal of Virology*, 77(10):5721–5730.

Friedman, J., Cho, W.-K., Chu, C. K., Keedy, K. S., Archin, N. M., Margolis, D. M., and Karn, J. (2011). Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of Zeste 2. *Journal of Virology*, 85(17):9078–9089.

Fuda, N. J., Ardehali, M. B., and Lis, J. T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature*, 461(7261):186–192.

Fujinaga, K., Barboric, M., Li, Q., Luo, Z., Price, D. H., and Peterlin, B. M. (2012). PKC phosphorylates HEXIM1 and regulates P-TEFb activity. *Nucleic Acids Research*, 40(18):9160–9170.

Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D., Sidhu, G. D., Stahl, R. E., Zolla-Pazner, S., Leibowitch, J., and Popovic, M. (1983). Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science*, 220(4599):865–867.

Gandhi, R. T., Coombs, R. W., Chan, E. S., Bosch, R. J., Zheng, L., Margolis, D. M., Read, S., Kallungal, B., Chang, M., Goecker, E. A., Wiegand, A., Kearney, M., Jacobson, J. M., D'Aquila, R., Lederman, M. M., Mellors, J. W., Eron, J. J., and AIDS Clinical Trials Group (ACTG) A5244 Team (2012). No effect of raltegravir intensification on viral replication markers in the blood of HIV-1-infected patients receiving antiretroviral therapy. *Journal of acquired immune deficiency syndromes (1999)*, 59(3):229–235.

Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., Shaw, G. M., Sharp, P. M., and Hahn, B. H. (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature*, 397(6718):436–441.

Gaspar, H. B., Parsley, K. L., Howe, S., King, D., Gilmour, K. C., Sinclair, J., Brouns, G., Schmidt, M., von Kalle, C., Barington, T., Jakobsen, M. A., Christensen, H. O., Al Ghonaium, A., White, H. N., Smith, J. L., Levinsky, R. J., Ali, R. R., Kinnon, C., and Thrasher, A. J. (2004). Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet (London, England)*, 364(9452):2181–2187.

Ge, H., Si, Y., and Roeder, R. G. (1998). Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *The EMBO journal*, 17(22):6723–6729.

Gérard, A., Ségéral, E., Naughtin, M., Abdouni, A., Charmeteau, B., Cheynier, R., Rain, J.-C., and Emiliani, S. (2015). The Integrase Cofactor LEDGF/p75 Associates with Iws1 and Spt6 for Postintegration Silencing of HIV-1 Gene Expression in Latently Infected Cells. *Cell Host and Microbe*, 17(1):107–117.

Gijsbers, R., Ronen, K., Vets, S., Malani, N., De Rijck, J., McNeely, M., Bushman, F. D., and Debyser, Z. (2009). LEDGF Hybrids Efficiently Retarget Lentiviral Integration Into Heterochromatin. *Molecular Therapy*, 18(3):552–560.

Gijsbers, R., Vets, S., De Rijck, J., Ocwieja, K. E., Ronen, K., Malani, N., Bushman, F. D., and Debyser, Z. (2011a). Role of the PWWP domain of lens epithelium-derived growth factor (LEDGF)/p75 cofactor in lentiviral integration targeting. *The Journal of biological chemistry*, 286(48):41812–41825.

Gijsbers, R., Vets, S., De Rijck, J., Ocwieja, K. E., Ronen, K., Malani, N., Bushman, F. D., and Debyser, Z. (2011b). Role of the PWWP Domain of Lens Epithelium-derived Growth Factor (LEDGF)/p75 Cofactor in Lentiviral Integration Targeting. *Journal of Biological Chemistry*, 286(48):41812–41825.

Ginn, S. L., Liao, S. H. Y., Dane, A. P., Hu, M., Hyman, J., Finnie, J. W., Zheng, M., Cavazzana-Calvo, M., Alexander, S. I., Thrasher, A. J., and Alexander, I. E. (2010). Lymphomagenesis in SCID-X1 mice following lentivirus-mediated phenotype correction independent of insertional mutagenesis and gammac overexpression. *Molecular therapy : the journal of the American Society of Gene Therapy*, 18(5):965–976.

Glass, W. G., McDermott, D. H., Lim, J. K., Lekhong, S., Yu, S. F., Frank, W. A., Pape, J., Cheshier, R. C., and Murphy, P. M. (2006). CCR5 deficiency increases risk of symptomatic West Nile virus infection. *The Journal of experimental medicine*, 203(1):35–40.

Goonetilleke, N., Liu, M. K. P., Salazar-Gonzalez, J. F., Ferrari, G., Giorgi, E., Ganusov, V. V., Keele, B. F., Learn, G. H., Turnbull, E. L., Salazar, M. G., Weinhold, K. J., Moore, S., CHAVI Clinical Core B, Letvin, N., Haynes, B. F., Cohen, M. S., Hraber, P., Bhattacharya, T., Borrow, P., Perelson, A. S., Hahn, B. H., Shaw, G. M., Korber, B. T., and McMichael, A. J. (2009). The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *The Journal of experimental medicine*, 206(6):1253–1272.

Grez, M., Reichenbach, J., Schwäble, J., Seger, R., Dinauer, M. C., and Thrasher, A. J. (2011). Gene therapy of chronic granulomatous disease: the engraftment dilemma. *Molecular therapy : the journal of the American Society of Gene Therapy*, 19(1):28–35.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, 453(7197):948–951.

Gupta, K., Brady, T., Dyer, B. M., Malani, N., Hwang, Y., Male, F., Nolte, R. T., Wang, L., Velthuisen, E., Jeffrey, J., Van Duyne, G. D., and Bushman, F. D. (2014). Allosteric inhibition of human immunodeficiency virus integrase: late block during viral replication and abnormal multimerization involving specific protein domains. *The Journal of biological chemistry*, 289(30):20477–20488.

Gupta, S. S., Maetzig, T., Maertens, G. N., Sharif, A., Rothe, M., Weidner-Glunde, M., Galla, M., Schambach, A., Cherepanov, P., and Schulz, T. F. (2013). Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *Journal of Virology*, 87(23):12721–12736.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B. H., Wintergerst, U., Velez, M. C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F. D., Fischer, A.,

and Cavazzana-Calvo, M. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *Journal of Clinical Investigation*, 118(9):3132–3142.

Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.-L., Fraser, C. C., Cavazzana-Calvo, M., and Fischer, A. (2003a). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *The New England journal of medicine*, 348(3):255–256.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J. I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L. E., Wissler, M., Prinz, C., Rabbitts, T. H., Le Deist, F., Fischer, A., and Cavazzana-Calvo, M. (2003b). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, 302(5644):415–419.

Han, Y., Lassen, K., Monie, D., Sedaghat, A. R., Shimoji, S., Liu, X., Pierson, T. C., Margolick, J. B., Siliciano, R. F., and Siliciano, J. D. (2004). Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *Journal of Virology*, 78(12):6122–6133.

Han, Y., Lin, Y. B., An, W., Xu, J., Yang, H.-C., O'Connell, K., Dordai, D., Boeke, J. D., Siliciano, J. D., and Siliciano, R. F. (2008). Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host and Microbe*, 4(2):134–146.

Hare, S., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature*, 464(7286):232–236.

Hare, S., Maertens, G. N., and Cherepanov, P. (2012). 3'-processing and strand transfer catalysed by retroviral integrase in crystallo. *The EMBO journal*, 31(13):3020–3028.

Hare, S., Shun, M.-C., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2009). A Novel Co-Crystal Structure Affords the Design of Gain-of-Function Lentiviral Integrase Mutants in the Presence of Modified PSIP1/LEDGF/p75. *PLoS Pathogens*, 5(1):e1000259.

Hatano, H., Strain, M. C., Scherzer, R., Bacchetti, P., Wentworth, D., Hoh, R., Martin, J. N., McCune, J. M., Neaton, J. D., Tracy, R. P., Hsue, P. Y., Richman, D. D., and Deeks, S. G. (2013). Increase in 2-long terminal repeat circles and decrease in D-dimer after raltegravir intensification in patients with treated HIV infection: a randomized, placebo-controlled trial. *The Journal of infectious diseases*, 208(9):1436–1442.

Hayouka, Z., Rosenbluh, J., Levin, A., Loya, S., Lebendiker, M., Veprintsev, D., Kotler, M., Hizi, A., Loyter, A., and Friedler, A. (2007). Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. *Proceedings of the National Academy of Sciences*, 104(20):8316–8321.

Hendrix, J., Gijsbers, R., De Rijck, J., Voet, A., Hotta, J.-i., McNeely, M., Hofkens, J., Debyser, Z., and Engelborghs, Y. (2011). The transcriptional co-activator LEDGF/p75 displays a dynamic scan-and-lock mechanism for chromatin tethering. *Nucleic Acids Research*, 39(4):1310–1325.

Henrich, T. J., Hanhauser, E., Marty, F. M., Sirignano, M. N., Keating, S., Lee, T.-H., Robles, Y. P., Davis, B. T., Li, J. Z., Heisey, A., Hill, A. L., Busch, M. P., Armand, P., Soiffer, R. J., Altfeld, M., and Kuritzkes, D. R. (2014). Antiretroviral-free HIV-1 remission and viral rebound after allogeneic stem cell transplantation: report of 2 cases. *Annals of internal medicine*, 161(5):319–327.

Ho, Y.-C., Shan, L., Hosmane, N. N., Wang, J., Laskey, S. B., Rosenbloom, D. I. S., Lai, J., Blankson, J. N., Siliciano, J. D., and Siliciano, R. F. (2013). Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. *Cell*, 155(3):540–551.

Hocqueloux, L., Avettand-Fènoël, V., Jacquot, S., Prazuck, T., Legac, E., Melard, A., Niang, M., Mille, C., Le Moal, G., Viard, J.-P., Rouzioux, C., and AC32 (Coordinated Action on HIV Reservoirs) of the Agence Nationale de Recherches sur le Sida et les Hépatites Virales (ANRS) (2013). Long-term antiretroviral therapy initiated during primary HIV-1 infection is key to achieving both low HIV reservoirs and normal T cell counts. *The Journal of antimicrobial chemotherapy*, 68(5):1169–1178.

Hocum, J. D., Battrell, L. R., Maynard, R., Adair, J. E., Beard, B. C., Rawlings, D. J., Kiem, H.-P., Miller, D. G., and Trobridge, G. D. (2015). VISA–Vector Integration Site Analysis server: a web-based server to rapidly identify retroviral integration sites from next-generation sequencing. *BMC Bioinformatics*, 16(1):212.

Hoen, B., Cooper, D. A., Lampe, F. C., Perrin, L., Clumeck, N., Phillips, A. N., Goh, L.-E., Lindback, S., Sereni, D., Gazzard, B., Montaner, J., Stellbrink, H.-J., Lazzarin, A., Ponscarme, D., Staszewski, S., Mathiesen, L., Smith, D., Finlayson, R., Weber, R., Wegmann, L., Janossy, G., Kinloch-de Loes, S., and QUEST Study Group (2007). Predictors of virological outcome and safety in primary HIV type 1-infected patients initiating quadruple antiretroviral therapy: QUEST GW PROB3005. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 45(3):381–390.

Holman, A. G. and Coffin, J. M. (2005). Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *Proceedings of the National Academy of Sciences of the United States of America*, 102(17):6103–6107.

Holmes-Son, M. (2002). Correct Integration Mediated by Integrase–LexA Fusion Proteins Incorporated into HIV-1. *Molecular Therapy*, 5(4):360–370.

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G. M., Kohn, D. B., Gregory, P. D., Holmes, M. C., and Cannon, P. M. (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nature Biotechnology*, 28(8):839–847.

Hoque, M., Shamanna, R. A., Guan, D., Pe'ery, T., and Mathews, M. B. (2011). HIV-1 replication and latency are regulated by translational control of cyclin T1. *Journal of Molecular Biology*, 410(5):917–932.

Howe, S. J., Mansour, M. R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M. H., Pike-Overzet, K., Chatters, S. J., de Ridder, D., Gilmour, K. C., Adams, S., Thornhill, S. I., Parsley, K. L., Staal, F. J. T., Gale, R. E., Linch, D. C., Bayford, J., Brown, L., Quaye, M., Kinnon, C., Ancliff, P., Webb, D. K., Schmidt, M., von Kalle, C., Gaspar, H. B., and Thrasher, A. J. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *Journal of Clinical Investigation*, 118(9):3143–3150.

Hughes, S., Jenkins, V., Dar, M. J., Engelman, A., and Cherepanov, P. (2010). Transcriptional co-activator LEDGF interacts with Cdc7-activator of S-phase kinase (ASK) and stimulates its enzymatic activity. *The Journal of biological chemistry*, 285(1):541–554.

Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müssig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O., Blau, I. W., Hofmann, W. K., and Thiel, E. (2009). Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *The New England journal of medicine*, 360(7):692–698.

Hymes, K. B., Cheung, T., Greene, J. B., Prose, N. S., Marcus, A., Ballard, H., William, D. C., and Laubenstein, L. J. (1981). Kaposi's sarcoma in homosexual men-a report of eight cases. *Lancet (London, England)*, 2(8247):598–600.

Ibrahimi, A., Vande Velde, G., Reumers, V., Toelen, J., Thiry, I., Vandeputte, C., Vets, S., Deroose, C., Bormans, G., Baekelandt, V., Debyser, Z., and Gijsbers, R. (2009). Highly efficient multicistronic lentiviral vectors with peptide 2A sequences. *Human Gene Therapy*, 20(8):845–860.

Irion, S., Luche, H., Gadue, P., Fehling, H. J., Kennedy, M., and Keller, G. (2007). Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nature Biotechnology*, 25(12):1477–1482.

Izumoto, Y., Kuroda, T., Harada, H., Kishimoto, T., and Nakamura, H. (1997). Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. *Biochemical and biophysical research communications*, 238(1):26–32.

Jaskolski, M., Alexandratos, J. N., Bujacz, G., and Wlodawer, A. (2009). Piecing together the structure of retroviral integrase, an important target in AIDS therapy. *The FEBS journal*, 276(11):2926–2946.

Jern, P. and Coffin, J. M. (2008). Effects of Retroviruses on Host Genome Function. *Annual Review of Genetics*, 42(1):709–732.

Jin, H., Li, D., Sivakumaran, H., Lor, M., Rustanti, L., Cloonan, N., Wani, S., and Harrich, D. (2016). Shutdown of HIV-1 Transcription in T Cells by Nullbasic, a Mutant Tat Protein. *mBio*, 7(4):e00518–16.

Jones, K. A. (1997). Taking a new TAK on tat transactivation. *Genes & development*, 11(20):2593–2599.

Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *The EMBO journal*, 22(8):1868–1877.

Jordan, A., Defechereux, P., and Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *The EMBO journal*, 20(7):1726–1738.

Jurado, K. A., Wang, H., Slaughter, A., Feng, L., Kessl, J. J., Koh, Y., Wang, W., Ballandras-Colas, A., Patel, P. A., Fuchs, J. R., Kvaratskhelia, M., and Engelman, A. (2013). Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(21):8690–8695.

Kauder, S. E., Bosque, A., Lindqvist, A., Planelles, V., and Verdin, E. (2009). Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathogens*, 5(6):e1000495.

Kaufmann, K. B., Brendel, C., Suerth, J. D., Mueller-Kuller, U., Chen-Wichmann, L., Schwäble, J., Pahujani, S., Kunkel, H., Schambach, A., Baum, C., and Grez, M. (2013). Alpharetroviral vector-mediated gene therapy for X-CGD: functional correction and lack of aberrant splicing. *Molecular therapy : the journal of the American Society of Gene Therapy*, 21(3):648–661.

Kearney, M. F., Spindler, J., Shao, W., Yu, S., Anderson, E. M., O'Shea, A., Rehm, C., Poethke, C., Kovacs, N., Mellors, J. W., Coffin, J. M., and Maldarelli, F. (2014). Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. *PLoS Pathogens*, 10(3):e1004010.

Keedy, K. S., Archin, N. M., Gates, A. T., Espeseth, A., Hazuda, D. J., and Margolis, D. M. (2009). A limited group of class I histone deacetylases acts to repress human immunodeficiency virus type 1 expression. *Journal of Virology*, 83(10):4749–4756.

Kessl, J. J., Jena, N., Koh, Y., Taskent-Sezgin, H., Slaughter, A., Feng, L., de Silva, S., Wu, L., Le Grice, S. F. J., Engelman, A., Fuchs, J. R., and Kvaratskhelia, M. (2012). Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *The Journal of biological chemistry*, 287(20):16801–16811.

Kinsella, R. J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., Kersey, P., and Flicek, P. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database : the journal of biological databases and curation*, 2011(0):bar030–bar030.

Kordelas, L., Verheyen, J., Beelen, D. W., Horn, P. A., Heinold, A., Kaiser, R., Trenschel, R., Schadendorf, D., Dittmer, U., Esser, S., and Essen HIV AlloSCT Group (2014). Shift of HIV tropism in stem-cell transplantation with CCR5 Delta32 mutation. *The New England journal of medicine*, 371(9):880–882.

Kotin, R. M., Linden, R. M., and Berns, K. I. (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *The EMBO journal*, 11(13):5071–5078.

Kustikova, O. S., Schiedlmeier, B., Brugman, M. H., Stahlhut, M., Bartels, S., Li, Z., and Baum, C. (2009). Cell-intrinsic and vector-related properties cooperate to determine the incidence and consequences of insertional mutagenesis. *Molecular therapy : the journal of the American Society of Gene Therapy*, 17(9):1537–1547.

LaFave, M. C., Varshney, G. K., Gildea, D. E., Wolfsberg, T. G., Baxevanis, A. D., and Burgess, S. M. (2014). MLV integration site selection is driven by strong enhancers and active promoters. *Nucleic Acids Research*, 42(7):4257–4269.

Laird, G. M., Bullen, C. K., Rosenbloom, D. I. S., Martin, A. R., Hill, A. L., Durand, C. M., Siliciano, J. D., and Siliciano, R. F. (2015). Ex vivo analysis identifies effective HIV-1 latency-reversing drug combinations. *Journal of Clinical Investigation*, 125(5):1901–1912.

Larena, M., Regner, M., and Lobigs, M. (2012). The chemokine receptor CCR5, a therapeutic target for HIV/AIDS antagonists, is critical for recovery in a mouse model of Japanese encephalitis. *PLoS ONE*, 7(9):e44834.

Le, T., Wright, E. J., Smith, D. M., He, W., Catano, G., Okulicz, J. F., Young, J. A., Clark, R. A., Richman, D. D., Little, S. J., and Ahuja, S. K. (2013). Enhanced CD4+ T-cell recovery with earlier HIV-1 antiretroviral therapy. *The New England journal of medicine*, 368(3):218–230.

Le Rouzic, E., Bonnard, D., Chasset, S., Bruneau, J.-M., Chevreuil, F., Le Strat, F., Nguyen, J., Beauvoir, R., Amadori, C., Brias, J., Vomscheid, S., Eiler, S., Lévy, N., Delelis, O., Deprez, E., Saïb, A., Zamborlini, A., Emiliani, S., Ruff, M., Ledoussal, B., Moreau, F., and Benarous, R. (2013). Dual inhibition of HIV-1 replication by integrase-LEDGF allosteric inhibitors is predominant at the post-integration stage. *Retrovirology*, 10(1):144.

Lelek, M., Casartelli, N., Pellin, D., Rizzi, E., Souque, P., Severgnini, M., Di Serio, C., Fricke, T., Diaz-Griffero, F., Zimmer, C., Charneau, P., and Di Nunzio, F. (2015). Chromatin organization at the nuclear pore favours HIV replication. *Nature Communications*, 6:6483.

Lenasi, T., Contreras, X., and Peterlin, B. M. (2008). Transcriptional interference antagonizes proviral gene expression to promote HIV latency. *Cell Host and Microbe*, 4(2):123–133.

Lewin, S. R., Murray, J. M., Solomon, A., Wightman, F., Cameron, P. U., Purcell, D. J., Zaunders, J. J., Grey, P., Bloch, M., Smith, D., Cooper, D. A., and Kelleher, A. D. (2008). Virologic determinants of success after structured treatment interruptions of antiretrovirals in acute HIV-1 infection. *Journal of acquired immune deficiency syndromes (1999)*, 47(2):140–147.

Lewinski, M. K., Bisgrove, D., Shinn, P., Chen, H., Hoffmann, C., Hannenhalli, S., Verdin, E., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2005). Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *Journal of Virology*, 79(11):6610–6619.

Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2006). Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathogens*, 2(6):e60.

Li, L., Krymskaya, L., Wang, J., Henley, J., Rao, A., Cao, L.-F., Tran, C.-A., Torres-Coronado, M., Gardner, A., Gonzalez, N., Kim, K., Liu, P.-Q., Hofer, U., Lopez, E., Gregory, P. D., Liu, Q., Holmes, M. C., Cannon, P. M., Zaia, J. A., and DiGiusto, D. L. (2013). Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Molecular therapy : the journal of the American Society of Gene Therapy*, 21(6):1259–1269.

Li, X., Krishnan, L., Cherepanov, P., and Engelman, A. (2011). Structural biology of retroviral DNA integration. *Virology*, 411(2):194–205.

Liao, H.-K., Gu, Y., Diaz, A., Marlett, J., Takahashi, Y., Li, M., Suzuki, K., Xu, R., Hishida, T., Chang, C.-J., Esteban, C. R., Young, J., and Izpisua Belmonte, J. C. (2015). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nature Communications*, 6:6413.

Linette, G. P., Stadtmauer, E. A., Maus, M. V., Rapoport, A. P., Levine, B. L., Emery, L., Litzky, L., Bagg, A., Carreno, B. M., Cimino, P. J., Binder-Scholl, G. K., Smethurst, D. P., Gerry, A. B., Pumphrey, N. J., Bennett, A. D., Brewer, J. E., Dukes, J., Harper, J., Tayton-Martin, H. K., Jakobsen, B. K., Hassan, N. J., Kalos, M., and June, C. H. (2013). Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood*, 122(6):863–871.

Liu, H., Dow, E. C., Arora, R., Kimata, J. T., Bull, L. M., Arduino, R. C., and Rice, A. P. (2006). Integration of human immunodeficiency virus type 1 in untreated infection occurs preferentially within genes. *Journal of Virology*, 80(15):7765–7768.

Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhlmann, H., Koup, R. A., and Landau, N. R. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, 86(3):367–377.

Liu, R.-d., Wu, J., Shao, R., and Xue, Y.-h. (2014). Mechanism and factors that control HIV-1 transcription and latency activation. *Journal of Zhejiang University SCIENCE B*, 15(5):455–465.

Llano, M., Delgado, S., Vanegas, M., and Poeschla, E. M. (2004). Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *Journal of Biological Chemistry*, 279(53):55570–55577.

Llano, M., Vanegas, M., Hutchins, N., Thompson, D., Delgado, S., and Poeschla, E. M. (2006). Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *Journal of Molecular Biology*, 360(4):760–773.

Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D. F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., Damo, M., Pello, O. M., Holmes, M. C., Gregory, P. D., Gritti, A., Broccoli, V., Bonini, C., and Naldini, L. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nature Methods*, 8(10):861–869.

Lorenzo-Redondo, R., Fryer, H. R., Bedford, T., Kim, E.-Y., Archer, J., Pond, S. L. K., Chung, Y.-S., Penugonda, S., Chipman, J. G., Fletcher, C. V., Schacker, T. W., Malim, M. H., Rambaut, A., Haase, A. T., McLean, A. R., and Wolinsky, S. M. (2016). Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature*, 530(7588):51–56.

Maartens, G., Celum, C., and Lewin, S. R. (2014). HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet (London, England)*, 384(9939):258–271.

Maertens, G., Cherepanov, P., Debyser, Z., Engelborghs, Y., and Engelman, A. (2004). Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75. *Journal of Biological Chemistry*, 279(32):33421–33429.

Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *Journal of Biological Chemistry*, 278(35):33528–33539.

Maertens, G. N., Cherepanov, P., and Engelman, A. (2006). Transcriptional co-activator p75 binds and tethers the Myc-interacting protein JPO2 to chromatin. *Journal of Cell Science*, 119(Pt 12):2563–2571.

Maertens, G. N., Hare, S., and Cherepanov, P. (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature*, 468(7321):326–329.

Maetzig, T., Galla, M., Baum, C., and Schambach, A. (2011). Gammaretroviral vectors: biology, technology and application. 3(6):677–713.

Malatinkova, E., De Spiegelaere, W., Bonczkowski, P., Kiselinova, M., Vervisch, K., Trypsteen, W., Johnson, M., Verhofstede, C., de Looze, D., Murray, C., Kinloch-de Loes, S., and Vandekerckhove, L. (2015). Impact of a decade of successful antiretroviral therapy initiated at HIV-1 seroconversion on blood and rectal reservoirs. *eLife*, 4.

Maldarelli, F., Wu, X., Su, L., Simonetti, F. R., Shao, W., Hill, S., Spindler, J., Ferris, A. L., Mellors, J. W., Kearney, M. F., Coffin, J. M., and Hughes, S. H. (2014). HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*, 345(6193):179–183.

Marini, B., Kertesz-Farkas, A., Ali, H., Lucic, B., Lisek, K., Manganaro, L., Pongor, S., Luzzati, R., Recchia, A., Mavilio, F., Giacca, M., and Lusic, M. (2015). Nuclear architecture dictates HIV-1 integration site selection. *Nature*, 521(7551):227–231.

Marshall, H. M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E., and Bushman, F. D. (2007). Role of PSIP1/LEDGF/p75 in Lentiviral Infectivity and Integration Targeting. *PLoS ONE*, 2(12):e1340.

Mascola, J. R. and Haynes, B. F. (2013). HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunological reviews*, 254(1):225–244.

Maskell, D. P., Renault, L., Serrao, E., Lesbats, P., Matadeen, R., Hare, S., Lindemann, D., Engelman, A. N., Costa, A., and Cherepanov, P. (2015). Structural basis for retroviral integration into nucleosomes. *Nature*, 523(7560):366–369.

McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M., and Hope, T. J. (2002). Visualization of the intracellular behavior of HIV in living cells. *The Journal of cell biology*, 159(3):441–452.

McMichael, A. J., Borrow, P., Tomaras, G. D., Goonetilleke, N., and Haynes, B. F. (2009). The immune response during acute HIV-1 infection: clues for vaccine development. *Nature reviews. Immunology*, 10(1):11–23.

Metcalf Pate, K. A., Pohlmeyer, C. W., Walker-Sperling, V. E., Foote, J. B., Najarro, K. M., Cryer, C. G., Salgado, M., Gama, L., Engle, E. L., Shirk, E. N., Queen, S. E., Chioma, S., Vermillion, M. S., Bullock, B., Li, M., Lyons, C. E., Adams, R. J., Zink, M. C., Clements, J. E., Mankowski, J. L., and Blankson, J. N. (2015). A Murine Viral Outgrowth Assay to Detect Residual HIV Type 1 in Patients With Undetectable Viral Loads. *The Journal of infectious diseases*, 212(9):1387–1396.

Mitchell, R. S., Beitzel, B. F., Schroder, A. R. W., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2004). Retroviral DNA Integration: ASLV, HIV, and MLV Show Distinct Target Site Preferences. *PLoS Biology*, 2(8):e234.

Modlich, U., Bohne, J., Schmidt, M., Von Kalle, C., Knoss, S., Schambach, A., and Baum, C. (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood*, 108(8):2545–2553.

Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M. H., Schambach, A., Charrier, S., Galy, A., Thrasher, A. J., Bueren, J., and Baum, C. (2009a). Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Molecular therapy : the journal of the American Society of Gene Therapy*, 17(11):1919–1928.

Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M. H., Schambach, A., Charrier, S., Galy, A., Thrasher, A. J., Bueren, J., and Baum, C. (2009b). Insertional Transformation of Hematopoietic Cells by Self-inactivating Lentiviral and Gammaretroviral Vectors. *Molecular therapy : the journal of the American Society of Gene Therapy*, 17(11):1919–1928.

Moiani, A., Suerth, J. D., Gandolfi, F., Rizzi, E., Severgnini, M., De Bellis, G., Schambach, A., and Mavilio, F. (2014). Genome-wide analysis of alpharetroviral integration in human hematopoietic stem/progenitor cells. *Genes*, 5(2):415–429.

Moldt, B., Miskey, C., Staunstrup, N. H., Gogol-Döring, A., Bak, R. O., Sharma, N., Mátés, L., Izsvák, Z., Chen, W., Ivics, Z., and Mikkelsen, J. G. (2011). Comparative genomic integration profiling of Sleeping Beauty transposons mobilized with high efficacy from integrase-defective lentiviral vectors in primary human cells. *Molecular therapy : the journal of the American Society of Gene Therapy*, 19(8):1499–1510.

Montini, E. and Cesana, D. (2012). Genotoxicity assay for gene therapy vectors in tumor prone Cdkn2aâĄż/âĄż mice. *Methods in enzymology*, 507:171–185.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Bartholomae, C. C., Ranzani, M., Benedicenti, F., Sergi, L. S., Ambrosi, A., Ponzoni, M., Doglioni, C., Di Serio, C., von Kalle, C., and Naldini, L. (2009a). The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *Journal of Clinical Investigation*, 119(4):964–975.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Bartholomae, C. C., Ranzani, M., Benedicenti, F., Sergi, L. S., Ambrosi, A., Ponzoni, M., Doglioni, C., Di Serio, C., von Kalle, C., and Naldini, L. (2009b). The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *Journal of Clinical Investigation*, 119(4):964–975.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., Sergi, L. S., Benedicenti, F., Ambrosi, A., Di Serio, C., Doglioni, C., von Kalle, C., and Naldini, L. (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nature Biotechnology*, 24(6):687–696.

Mousseau, G., Kessing, C. F., Fromentin, R., Trautmann, L., Chomont, N., and Valente, S. T. (2015). The Tat Inhibitor Didehydro-Cortistatin A Prevents HIV-1 Reactivation from Latency. *mBio*, 6(4):e00465.

Mussolino, C. and Cathomen, T. (2013). RNA guides genome engineering. *Nature Biotechnology*, 31(3):208–209.

Naldini, L. (2015). Gene therapy returns to centre stage. *Nature*, 526(7573):351–360.

Newrzela, S., Cornils, K., Li, Z., Baum, C., Brugman, M. H., Hartmann, M., Meyer, J., Hartmann, S., Hansmann, M.-L., Fehse, B., and von Laer, D. (2008). Resistance of mature T cells to oncogene transformation. *Blood*, 112(6):2278–2286.

Nguyen, V. T., Kiss, T., Michels, A. A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, 414(6861):322–325.

Nowrouzi, A., Dittrich, M., Klanke, C., Heinkelein, M., Rammling, M., Dandekar, T., von Kalle, C., and Rethwilm, A. (2006). Genome-wide mapping of foamy virus vector integrations into a human cell line. *Journal of General Virology*, 87(Pt 5):1339–1347.

Ocwieja, K. E., Brady, T. L., Ronen, K., Huegel, A., Roth, S. L., Schaller, T., James, L. C., Towers, G. J., Young, J. A. T., Chanda, S. K., König, R., Malani, N., Berry, C. C., and Bushman, F. D. (2011). HIV Integration Targeting: A Pathway Involving Transportin-3 and the Nuclear Pore Protein RanBP2. *PLoS Pathogens*, 7(3):e1001313–14.

Onodera, M., Ariga, T., Kawamura, N., Kobayashi, I., Ohtsu, M., Yamada, M., Tame, A., Furuta, H., Okano, M., Matsumoto, S., Kotani, H., McGarrity, G. J., Blaese, R. M., and Sakiyama, Y. (1998). Successful peripheral T-lymphocyte-directed gene transfer for a patient with severe combined immune deficiency caused by adenosine deaminase deficiency. *Blood*, 91(1):30–36.

Ordovás, L., Boon, R., Pistoni, M., Chen, Y., Wolfs, E., Guo, W., Sambathkumar, R., Bobis-Wozowicz, S., Helsen, N., Vanhove, J., Berckmans, P., Cai, Q., Vanuytsel, K., Eggermont, K., Vanslembrouck, V., Schmidt, B. Z., Raitano, S., Van Den Bosch, L., Nahmias, Y., Cathomen, T., Struys, T., and Verfaillie, C. M. (2015). Efficient Recombinase-Mediated Cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals AAVS1 Locus-Mediated Transgene Inhibition. *Stem cell reports*, 5(5):918–931.

Osório, L., Gijsbers, R., Oliveras-Salvá, M., Michiels, A., Debyser, Z., Van den Haute, C., and Baekelandt, V. (2014). Viral vectors expressing a single microRNA-based short-hairpin RNA result in potent gene silencing in vitro and in vivo. *Journal of biotechnology*, 169:71–81.

Ott, M. G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kühlcke, K., Schilz, A., Kunkel, H., Naundorf, S., Brinkmann, A., Deichmann, A., Fischer, M., Ball, C., Pilz, I., Dunbar, C., Du, Y., Jenkins, N. A., Copeland, N. G., Lüthi, U., Hassan, M., Thrasher, A. J., Hoelzer, D., von Kalle, C., Seger, R., and Grez, M. (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nature Medicine*, 12(4):401–409.

Palacios, J. A., Pérez-Piñar, T., Toro, C., Sanz-Minguela, B., Moreno, V., Valencia, E., Gómez-Hernando, C., and Rodés, B. (2012). Long-term nonprogressor and elite controller patients who control viremia have a higher percentage of methylation in their HIV-1 proviral promoters than aviremic patients receiving highly active antiretroviral therapy. *Journal of Virology*, 86(23):13081–13084.

Palmer, S., Maldarelli, F., Wiegand, A., Bernstein, B., Hanna, G. J., Brun, S. C., Kempf, D. J., Mellors, J. W., Coffin, J. M., and King, M. S. (2008). Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proceedings of the National Academy of Sciences of the United States of America*, 105(10):3879–3884.

Pandeló José, D., Bartholomeeusen, K., da Cunha, R. D., Abreu, C. M., Glinski, J., da Costa, T. B. F., Bacchi Rabay, A. F. M., Pianowski Filho, L. F., Dudycz, L. W., Ranga, U., Peterlin, B. M., Pianowski, L. F., Tanuri, A., and Aguiar, R. S. (2014). Reactivation of latent HIV-1 by new semi-synthetic ingenol esters. *Virology*, 462-463:328–339.

Papapetrou, E. P., Lee, G., Malani, N., Setty, M., Riviere, I., Tirunagari, L. M. S., Kadota, K., Roth, S. L., Giardina, P., Viale, A., Leslie, C., Bushman, F. D., Studer, L., and Sadelain, M. (2011). Genomic safe harbors permit high β-globin transgene expression in thalassemia induced pluripotent stem cells. *Nature Biotechnology*, 29(1):73–78.

Papapetrou, E. P. and Schambach, A. (2016). Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Molecular therapy : the journal of the American Society of Gene Therapy*, 24(4):678–684.

Pascual-Garcia, P. and Capelson, M. (2014). Nuclear pores as versatile platforms for gene regulation. *Current opinion in genetics & development*, 25:110–117.

Perez, E. E., Wang, J., Miller, J. C., Jouvenot, Y., Kim, K. A., Liu, O., Wang, N., Lee, G., Bartsevich, V. V., Lee, Y.-L., Guschin, D. Y., Rupniewski, I., Waite, A. J., Carpenito, C., Carroll, R. G., Orange, J. S., Urnov, F. D., Rebar, E. J., Ando, D., Gregory, P. D., Riley, J. L., Holmes, M. C., and June, C. H. (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature Biotechnology*, 26(7):808–816.

Persaud, D., Gay, H., Ziemniak, C., Chen, Y. H., Piatak, M., Chun, T.-W., Strain, M., Richman, D., and Luzuriaga, K. (2013). Absence of detectable HIV-1 viremia after treatment cessation in an infant. *The New England journal of medicine*, 369(19):1828–1835.

Peterlin, B. M. and Price, D. H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Molecular Cell*, 23(3):297–305.

Pizzato, M., Erlwein, O., Bonsall, D., Kaye, S., Muir, D., and McClure, M. O. (2009). A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. *Journal of Virological Methods*, 156(1-2):1–7.

Pommier, Y., Johnson, A. A., and Marchand, C. (2005). Integrase inhibitors to treat HIV/Aids. *Nature Reviews Drug Discovery*, 4(3):236–248.

Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011). Atomic-level modelling of the HIV capsid. *Nature*, 469(7330):424–427.

Porter, D. L., Levine, B. L., Kalos, M., Bagg, A., and June, C. H. (2011). Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *The New England journal of medicine*, 365(8):725–733.

Pradeepa, M. M., Sutherland, H. G., Ule, J., Grimes, G. R., and Bickmore, W. A. (2012). Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS genetics*, 8(5):e1002717.

Pruss, D., Bushman, F. D., and Wolffe, A. P. (1994a). Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proceedings of the National Academy of Sciences*, 91(13):5913–5917.

Pruss, D., Reeves, R., Bushman, F. D., and Wolffe, A. P. (1994b). The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *Journal of Biological Chemistry*, 269(40):25031–25041.

Pryciak, P. M. and Varmus, H. E. (1992). Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell*, 69(5):769–780.

Quercioli, V., Di Primio, C., Casini, A., Mulder, L. C. F., Vranckx, L. S., Borrenberghs, D., Gijsbers, R., Debyser, Z., and Cereseto, A. (2016). Comparative Analysis of HIV-1 and Murine Leukemia Virus Three-Dimensional Nuclear Distributions. *Journal of Virology*, 90(10):5205–5209.

Rasmussen, T. A., Tolstrup, M., Brinkmann, C. R., Olesen, R., Erikstrup, C., Solomon, A., Winckelmann, A., Palmer, S., Dinarello, C., Buzon, M., Lichterfeld, M., Lewin, S. R., Østergaard, L., and Søgaard, O. S. (2014). Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *The lancet. HIV*, 1(1):e13–21.

Razooky, B. S., Pai, A., Aull, K., Rouzine, I. M., and Weinberger, L. S. (2015). A hardwired HIV latency program. *Cell*, 160(5):990–1001.

Richman, D. D., Margolis, D. M., Delaney, M., Greene, W. C., Hazuda, D., and Pomerantz, R. J. (2009). The challenge of finding a cure for HIV infection. *Science*, 323(5919):1304–1307.

Rohr, O., Marban, C., Aunis, D., and Schaeffer, E. (2003). Regulation of HIV-1 gene transcription: from lymphocytes to microglial cells. *Journal of leukocyte biology*, 74(5):736–749.

Rothe, M., Schambach, A., and Biasco, L. (2014a). Safety of Gene Therapy: New Insights to a Puzzling Case. *Current gene therapy*, 14(6):429–436.

Rothe, M., Schambach, A., and Biasco, L. (2014b). Safety of gene therapy: new insights to a puzzling case. *Current gene therapy*, 14(6):429–436.

Rouzine, I. M., Weinberger, A. D., and Weinberger, L. S. (2015). An evolutionary role for HIV latency in enhancing viral transmission. *Cell*, 160(5):1002–1012.

Sadelain, M., Papapetrou, E. P., and Bushman, F. D. (2011a). Safe harbours for the integration of new DNA in the human genome. *Nature Reviews Cancer*, 12(1):51–58.

Sadelain, M., Papapetrou, E. P., and Bushman, F. D. (2011b). Safe harbours for the integration of new DNA in the human genome. *Nature Reviews Cancer*, 12(1):51–58.

Sáez-Cirión, A., Bacchus, C., Hocqueloux, L., Avettand-Fènoël, V., Girault, I., Lecuroux, C., Potard, V., Versmisse, P., Melard, A., Prazuck, T., Descours, B., Guergnon, J., Viard, J.-P., Boufassa, F., Lambotte, O., Goujard, C., Meyer, L., Costagliola, D., Venet, A., Pancino, G., Autran, B., Rouzioux, C., and ANRS VISCONTI Study Group (2013). Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS Pathogens*, 9(3):e1003211.

Saleh, S., Wightman, F., Ramanayake, S., Alexander, M., Kumar, N., Khoury, G., Pereira, C., Purcell, D., Cameron, P. U., and Lewin, S. R. (2011). Expression and reactivation of HIV in a chemokine induced model of HIV latency in primary resting CD4+ T cells. *Retrovirology*, 8(1):80.

Sallusto, F., Lenig, D., Förster, R., Lipp, M., and Lanzavecchia, A. (1999). Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature*, 401(6754):708–712.

Schrijvers, R., De Rijck, J., Demeulemeester, J., Adachi, N., Vets, S., Ronen, K., Christ, F., Bushman, F. D., Debyser, Z., and Gijsbers, R. (2012a). LEDGF/p75-Independent HIV-1 Replication Demonstrates a Role for HRP-2 and Remains Sensitive to Inhibition by LEDGINs. *PLoS Pathogens*, 8(3):e1002558.

Schrijvers, R., Vets, S., De Rijck, J., Malani, N., Bushman, F. D., Debyser, Z., and Gijsbers, R. (2012b). HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology*, 9(1):84.

Schroder, A. R. W., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, 110(4):521–529.

Schrödinger, LLC (2015). The PyMOL molecular graphics system, version 1.8.

Sekhar, V. and McBride, A. A. (2012). Phosphorylation regulates binding of the human papillomavirus type 8 E2 protein to host chromosomes. *Journal of Virology*, 86(18):10047–10058.

Sekhar, V., Reed, S. C., and McBride, A. A. (2010). Interaction of the betapapillomavirus E2 tethering protein with mitotic chromosomes. *Journal of Virology*, 84(1):543–557.

Serrao, E., Ballandras-Colas, A., Cherepanov, P., Maertens, G. N., and Engelman, A. N. (2015). Key determinants of target DNA recognition by retroviral intasomes. *Retrovirology*, 12(1):39.

Serrao, E., Krishnan, L., Shun, M.-C., Li, X., Cherepanov, P., Engelman, A., and Maertens, G. N. (2014). Integrase residues that determine nucleotide preferences at sites of HIV-1 integration: implications for the mechanism of target DNA binding. *Nucleic Acids Research*, 42(8):5164–5176.

Shan, L., Deng, K., Shroff, N. S., Durand, C. M., Rabi, S. A., Yang, H.-C., Zhang, H., Margolick, J. B., Blankson, J. N., and Siliciano, R. F. (2012). Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. *Immunity*, 36(3):491–501.

Shan, L., Yang, H.-C., Rabi, S. A., Bravo, H. C., Shroff, N. S., Irizarry, R. A., Zhang, H., Margolick, J. B., Siliciano, J. D., and Siliciano, R. F. (2011). Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. *Journal of Virology*, 85(11):5384–5393.

Shang, H.-t., Ding, J.-w., Yu, S.-y., Wu, T., Zhang, Q.-l., and Liang, F.-j. (2015). Progress and challenges in the use of latent HIV-1 reactivating agents. *Acta pharmacologica Sinica*, 36(8):908–916.

Sharma, A., Larue, R. C., Plumb, M. R., Malani, N., Male, F., Slaughter, A., Kessl, J. J., Shkriabai, N., Coward, E., Aiyer, S. S., Green, P. L., Wu, L., Roth, M. J., Bushman, F. D., and Kvaratskhelia, M. (2013). BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proceedings of the National Academy of Sciences of the United States of America*, 110(29):12036–12041.

Sharma, A., Slaughter, A., Jena, N., Feng, L., Kessl, J. J., Fadel, H. J., Malani, N., Male, F., Wu, L., Poeschla, E., Bushman, F. D., Fuchs, J. R., and Kvaratskhelia, M. (2014). A New Class of Multimerization Selective Inhibitors of HIV-1 Integrase. *PLoS Pathogens*, 10(5):e1004171.

Sharp, P. M., Bailes, E., Chaudhuri, R. R., Rodenburg, C. M., Santiago, M. O., and Hahn, B. H. (2001). The origins of acquired immune deficiency syndrome viruses: where and when? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 356(1410):867–876.

Sherrill-Mix, S., Lewinski, M. K., Famiglietti, M., Bosque, A., Malani, N., Ocwieja, K. E., Berry, C. C., Looney, D., Shan, L., Agosto, L. M., Pace, M. J., Siliciano, R. F., O'Doherty, U., Guatelli, J., Planelles, V., and Bushman, F. D. (2013). HIV latency and integration site placement in five cell-based models. *Retrovirology*, 10(1):90.

Shinohara, T., Singh, D. P., and Fatma, N. (2002). LEDGF, a survival factor, activates stress-related genes. *Progress in retinal and eye research*, 21(3):341–358.

Shiu, C., Cunningham, C. K., Greenough, T., Muresan, P., Sanchez-Merino, V., Carey, V., Jackson, J. B., Ziemniak, C., Fox, L., Belzer, M., Ray, S. C., Luzuriaga, K., Persaud, D., and Pediatric AIDS Clinical Trials Group P1059 Team (2009). Identification of ongoing human immunodeficiency virus type 1 (HIV-1) replication in residual viremia during recombinant HIV-1 poxvirus immunizations in patients with clinically undetectable viral loads on durable suppressive highly active antiretroviral therapy. *Journal of Virology*, 83(19):9731–9742.

Shkriabai, N., Dharmarajan, V., Slaughter, A., Kessl, J. J., Larue, R. C., Feng, L., Fuchs, J. R., Griffin, P. R., and Kvaratskhelia, M. (2014). A critical role of the C-terminal segment for allosteric inhibitor-induced aberrant multimerization of HIV-1 integrase. *The Journal of biological chemistry*, 289(38):26430–26440.

Shou, Y., Ma, Z., Lu, T., and Sorrentino, B. P. (2006). Unique risk factors for insertional mutagenesis in a mouse model of XSCID gene therapy. *Proceedings of the National Academy of Sciences*, 103(31):11730–11735.

Shun, M.-C., Raghavendra, N. K., Vandegraaff, N., Daigle, J. E., Hughes, S., Kellam, P., Cherepanov, P., and Engelman, A. (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes & development*, 21(14):1767–1778.

Siliciano, J. D., Kajdas, J., Finzi, D., Quinn, T. C., Chadwick, K., Margolick, J. B., Kovacs, C., Gange, S. J., and Siliciano, R. F. (2003). Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nature Medicine*, 9(6):727–728.

Siliciano, J. D. and Siliciano, R. F. (2004). A long-term latent reservoir for HIV-1: discovery and clinical implications. *The Journal of antimicrobial chemotherapy*, 54(1):6–9.

Siliciano, J. M. and Siliciano, R. F. (2015). The Remarkable Stability of the Latent Reservoir for HIV-1 in Resting Memory CD4+ T Cells. *The Journal of infectious diseases*.

Silvers, R. M., Smith, J. A., Schowalter, M., Litwin, S., Liang, Z., Geary, K., and Daniel, R. (2010). Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein. *Human Gene Therapy*, 21(3):337–349.

Simon, V., Bloch, N., and Landau, N. R. (2015). Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nature immunology*, 16(6):546–553.

Singh, A., Razooky, B., Cox, C. D., Simpson, M. L., and Weinberger, L. S. (2010). Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical journal*, 98(8):L32–4.

Søgaard, O. S., Graversen, M. E., Leth, S., Olesen, R., Brinkmann, C. R., Nissen, S. K., Kjaer, A. S., Schleimann, M. H., Denton, P. W., Hey-Cunningham, W. J., Koelsch, K. K., Pantaleo, G., Krogsgaard, K., Sommerfelt, M., Fromentin, R., Chomont, N., Rasmussen, T. A., Østergaard, L., and Tolstrup, M. (2015). The Depsipeptide Romidepsin Reverses HIV-1 Latency In Vivo. *PLoS Pathogens*, 11(9):e1005142.

Sood, V. and Brickner, J. H. (2014). Nuclear pore interactions with the genome. *Current opinion in genetics & development*, 25:43–49.

Spina, C. A., Anderson, J., Archin, N. M., Bosque, A., Chan, J., Famiglietti, M., Greene, W. C., Kashuba, A., Lewin, S. R., Margolis, D. M., Mau, M., Ruelas, D., Saleh, S., Shirakawa, K., Siliciano, R. F., Singhania, A., Soto, P. C., Terry, V. H., Verdin, E., Woelk, C., Wooden, S., Xing, S., and Planelles, V. (2013). An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. *PLoS Pathogens*, 9(12):e1003834.

Staunstrup, N. H., Moldt, B., Mátés, L., Villesen, P., Jakobsen, M., Ivics, Z., Izsvák, Z., and Mikkelsen, J. G. (2009). Hybrid Lentivirus-transposon Vectors With a Random Integration Profile in Human Cells. *Nature*, 17(7):1205–1214.

Stec, I., Nagl, S. B., van Ommen, G. J., and den Dunnen, J. T. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS letters*, 473(1):1–5.

Stein, S., Ott, M. G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., Schmidt, M., Krämer, A., Schwäble, J., Glimm, H., Koehl, U., Preiss, C., Ball, C., Martin, H., Göhring, G., Schwarzwaelder, K., Hofmann, W.-K., Karakaya, K., Tchatchou, S., Yang, R., Reinecke, P., Kühlcke, K., Schlegelberger, B., Thrasher, A. J., Hoelzer, D., Seger, R., von Kalle, C., and Grez, M. (2010). Genomic

instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nature Medicine*, 16(2):198–204.

Stein, S., Scholz, S., Schwäble, J., Sadat, M. A., Modlich, U., Schultze-Strasser, S., Diaz, M., Chen-Wichmann, L., Müller-Kuller, U., Brendel, C., Fronza, R., Kaufmann, K. B., Naundorf, S., Pech, N. K., Travers, J. B., Matute, J. D., Presson, R. G., Sandusky, G. E., Kunkel, H., Rudolf, E., Dillmann, A., von Kalle, C., Kühlcke, K., Baum, C., Schambach, A., Dinauer, M. C., Schmidt, M., and Grez, M. (2013). From bench to bedside: preclinical evaluation of a self-inactivating gammaretroviral vector for the gene therapy of X-linked chronic granulomatous disease. *Human gene therapy. Clinical development*, 24(2):86–98.

Stella, S. and Montoya, G. (2015). The genome editing revolution: A CRISPR-Cas TALE off-target story. *Inside the Cell*, 1(1):7–16.

Stevens, S. W. and Griffith, J. D. (1996). Sequence analysis of the human DNA flanking sites of human immunodeficiency virus type 1 integration. *Journal of Virology*, 70(9):6459–6462.

Suerth, J. D., Maetzig, T., Brugman, M. H., Heinz, N., Appelt, J.-U., Kaufmann, K. B., Schmidt, M., Grez, M., Modlich, U., Baum, C., and Schambach, A. (2012). Alpharetroviral self-inactivating vectors: long-term transgene expression in murine hematopoietic cells and low genotoxicity. *Molecular therapy : the journal of the American Society of Gene Therapy*, 20(5):1022–1032.

Sung, T.-L. and Rice, A. P. (2006). Effects of prostratin on Cyclin T1/P-TEFb function and the gene expression profile in primary resting CD4+ T cells. *Retrovirology*, 3(1):66.

Swiggard, W. J., Baytop, C., Yu, J. J., Dai, J., Li, C., Schretzenmair, R., Theodosopoulos, T., and O'Doherty, U. (2005). Human immunodeficiency virus type 1 can establish latent infection in resting CD4+ T cells in the absence of activating stimuli. *Journal of Virology*, 79(22):14179–14188.

Tabernilla, A. and Poveda, E. (2015). The START Trial: Definitive Evidence to Treat All HIV-Positive Persons Regardless of CD4 Counts. *AIDS reviews*, 17(3):187.

Tan, W., Dong, Z., Wilkinson, T. A., Barbas, C. F., and Chow, S. A. (2006). Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *Journal of Virology*, 80(4):1939–1948.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Structural & Molecular Biology*, 14(11):1025–1040.

Tebas, P., Stein, D., Tang, W. W., Frank, I., Wang, S. Q., Lee, G., Spratt, S. K., Surosky, R. T., Giedlin, M. A., Nichol, G., Holmes, M. C., Gregory, P. D., Ando, D. G., Kalos, M., Collman, R. G., Binder-Scholl, G., Plesa, G., Hwang, W.-T., Levine, B. L., and June, C. H. (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *The New England journal of medicine*, 370(10):901–910.

The INSIGHT START Study Group (2015). Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *The New England journal of medicine*, 373(9):795–807.

Tobaly-Tapiero, J., Bittoun, P., Lehmann-Che, J., Delelis, O., Giron, M.-L., de Thé, H., and Saïb, A. (2008). Chromatin Tethering of Incoming Foamy Virus by the Structural Gag Protein. *Traffic*, 9(10):1717–1727.

Torres, R. A. and Lewis, W. (2014). Aging and HIV/AIDS: pathogenetic role of therapeutic side effects. *Laboratory investigation; a journal of technical methods and pathology*, 94(2):120–128.

Touzot, F., Moshous, D., Creidy, R., Neven, B., Frange, P., Cros, G., Caccavelli, L., Blondeau, J., Magnani, A., Luby, J.-M., Ternaux, B., Picard, C., Blanche, S., Fischer, A., Hacein-Bey-Abina, S., and Cavazzana, M. (2015). Faster T-cell development following gene therapy compared with haploidentical HSCT in the treatment of SCID-X1. *Blood*, 125(23):3563–3569.

Trobridge, G. D., Miller, D. G., Jacobs, M. A., Allen, J. M., Kiem, H.-P., Kaul, R., and Russell, D. W. (2006). Foamy virus vector integration sites in normal human cells. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5):1498–1503.

Trono, D. (2012). Gene therapy: too much splice can spoil the dish. *Journal of Clinical Investigation*, 122(5):1600–1602.

Trono, D., Van Lint, C., Rouzioux, C., Verdin, E., Barré-Sinoussi, F., Chun, T.-W., and Chomont, N. (2010). HIV persistence and the prospect of long-term drug-free remissions for HIV-infected individuals. *Science*, 329(5988):174–180.

Trushin, S. A., Bren, G. D., Asin, S., Pennington, K. N., Paya, C. V., and Badley, A. D. (2005). Human immunodeficiency virus reactivation by phorbol esters or T-cell receptor ligation requires both PKCalpha and PKCtheta. *Journal of Virology*, 79(15):9821–9830.

Tsiang, M., Jones, G. S., Niedziela-Majka, A., Kan, E., Lansdon, E. B., Huang, W., Hung, M., Samuel, D., Novikov, N., Xu, Y., Mitchell, M., Guo, H., Babaoglu, K., Liu, X., Geleziunas, R., and Sakowicz, R. (2012). New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. *The Journal of biological chemistry*, 287(25):21189–21203.

Tsutsui, K. M., Sano, K., Hosoya, O., Miyamoto, T., and Tsutsui, K. (2011). Nuclear protein LEDGF/p75 recognizes supercoiled DNA by a novel DNA-binding domain. *Nucleic Acids Research*, 39(12):5067–5081.

Turchiano, G., Latella, M. C., Gogol-Döring, A., Cattoglio, C., Mavilio, F., Izsvák, Z., Ivics, Z., and Recchia, A. (2014). Genomic analysis of Sleeping Beauty transposon integration in human somatic cells. *PLoS ONE*, 9(11):e112712.

Turlure, F., Maertens, G., Rahman, S., Cherepanov, P., and Engelman, A. (2006). A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Research*, 34(5):1653–1665.

Van Lint, C., Emiliani, S., Ott, M., and Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *The EMBO journal*, 15(5):1112–1120.

van Nuland, R., van Schaik, F. M., Simonis, M., van Heesch, S., Cuppen, E., Boelens, R., Timmers, H. M., and van Ingen, H. (2013). Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. *Epigenetics & chromatin*, 6(1):12.

Vandekerckhove, L., Christ, F., Van Maele, B., De Rijck, J., Gijsbers, R., Van den Haute, C., Witvrouw, M., and Debyser, Z. (2006). Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *Journal of Virology*, 80(4):1886–1896.

Varadarajan, J., McWilliams, M. J., and Hughes, S. H. (2013). Treatment with suboptimal doses of raltegravir leads to aberrant HIV-1 integrations. *Proceedings of the National Academy of Sciences of the United States of America*, 110(36):14747–14752.

Verdin, E., Paras, P., and Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *The EMBO journal*, 12(8):3249–3259.

Verheyen, J., Esser, S., and Kordelas, L. (2014). More on shift of HIV tropism in stem-cell transplantation with CCR5 delta32/delta32 mutation. *The New England journal of medicine*, 371(25):2438–2438.

Vets, S., De Rijck, J., Brendel, C., Grez, M., Bushman, F., Debyser, Z., and Gijsbers, R. (2013). Transient Expression of an LEDGF/p75 Chimera Retargets Lentivector Integration and Functionally Rescues in a Model for X-CGD. *Molecular Therapy—Nucleic Acids*, 2(3):e77.

Võsa, L., Sudakov, A., Remm, M., Ustav, M., and Kurg, R. (2012). Identification and analysis of papillomavirus E2 protein binding sites in the human genome. *Journal of Virology*, 86(1):348–357.

Wagner, T. A., McLaughlin, S., Garg, K., Cheung, C. Y. K., Larsen, B. B., Styrchak, S., Huang, H. C., Edlefsen, P. T., Mullins, J. I., and Frenkel, L. M. (2014). HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science*, 345(6196):570–573.

Wainberg, M. A., Mesplède, T., and Quashie, P. K. (2012). The development of novel HIV integrase inhibitors and the problem of drug resistance. *Current opinion in virology*, 2(5):656–662.

Wang, H., Jurado, K. A., Wu, X., Shun, M.-C., Li, X., Ferris, A. L., Smith, S. J., Patel, P. A., Fuchs, J. R., Cherepanov, P., Kvaratskhelia, M., Hughes, S. H., and Engelman, A. (2012). HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. *Nucleic Acids Research*, 40(22):11518–11530.

Wang, H., Shun, M.-C., Li, X., Di Nunzio, F., Hare, S., Cherepanov, P., and Engelman, A. (2014). Efficient Transduction of LEDGF/p75 Mutant Cells by Gain-of-Function HIV-1 Integrase Mutant Viruses. *Molecular therapy. Methods & clinical development*, 1:2.

Weinberger, L. S., Burnett, J. C., Toettcher, J. E., Arkin, A. P., and Schaffer, D. V. (2005). Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, 122(2):169–182.

Weinberger, L. S., Dar, R. D., and Simpson, M. L. (2008). Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature genetics*, 40(4):466–470.

Whitney, J. B., Hill, A. L., Sanisetty, S., Penaloza-MacMaster, P., Liu, J., Shetty, M., Parenteau, L., Cabral, C., Shields, J., Blackmore, S., Smith, J. Y., Brinkman, A. L., Peter, L. E., Mathew, S. I., Smith, K. M., Borducchi, E. N., Rosenbloom, D. I. S., Lewis, M. G., Hattersley, J., Li, B., Hesselgesser, J., Geleziunas, R., Robb, M. L., Kim, J. H., Michael, N. L., and Barouch, D. H. (2014). Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature*, 512(7512):74–77.

Woods, N.-B., Bottero, V., Schmidt, M., von Kalle, C., and Verma, I. M. (2006). Gene therapy: therapeutic gene causing lymphoma. *Nature*, 440(7088):1123–1123.

Wu, X., Li, Y., Crise, B., and Burgess, S. M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science*, 300(5626):1749–1751.

Wu, X., Li, Y., Crise, B., Burgess, S. M., and Munroe, D. J. (2005). Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *Journal of Virology*, 79(8):5211–5214.

Wyl, V. v., Gianella, S., Fischer, M., Niederoest, B., Kuster, H., Battegay, M., Bernasconi, E., Cavassini, M., Rauch, A., Hirschel, B., Vernazza, P., Weber, R., Joos, B., Günthard, H. F., and Swiss HIV Cohort Study-SHCS (2011). Early antiretroviral therapy during primary HIV-1 infection results in a transient reduction of the viral setpoint upon treatment interruption. *PLoS ONE*, 6(11):e27463.

Yang, Z., Yik, J. H. N., Chen, R., He, N., Jang, M. K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Molecular Cell*, 19(4):535–545.

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, 414(6861):317–322.

Yin, Z., Shi, K., Banerjee, S., Pandey, K. K., Bera, S., Grandgenett, D. P., and Aihara, H. (2016). Crystal structure of the Rous sarcoma virus intasome. *Nature*, 530(7590):362–366.

Yokoyama, A. and Cleary, M. L. (2008). Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer cell*, 14(1):36–46.

Yukl, S. A., Boritz, E., Busch, M., Bentsen, C., Chun, T.-W., Douek, D., Eisele, E., Haase, A., Ho, Y.-C., Hütter, G., Justement, J. S., Keating, S., Lee, T.-H., Li, P., Murray, D., Palmer, S., Pilcher, C., Pillai, S., Price, R. W., Rothenberger, M., Schacker, T., Siliciano, J., Siliciano, R., Sinclair, E., Strain, M., Wong, J., Richman, D., and Deeks, S. G. (2013a). Challenges in detecting HIV persistence during potentially curative interventions: a study of the Berlin patient. *PLoS Pathogens*, 9(5):e1003347.

Yukl, S. A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A. L. M., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare, C. B., Pandori, M., Haase, A. T., Günthard, H. F., Fischer, M., Shergill, A. K., McQuaid, K., Havlir, D. V., and Wong, J. K. (2010). Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *The Journal of infectious diseases*, 202(10):1553–1561.

Yukl, S. A., Shergill, A. K., Ho, T., Killian, M., Girling, V., Epling, L., Li, P., Wong, L. K., Crouch, P., Deeks, S. G., Havlir, D. V., McQuaid, K., Sinclair, E., and Wong, J. K. (2013b). The distribution of HIV DNA and RNA in cell subsets differs in gut and blood of HIV-positive patients on ART: implications for viral persistence. *The Journal of infectious diseases*, 208(8):1212–1220.

Zhang, Z., Schuler, T., Zupancic, M., Wietgrefe, S., Staskus, K. A., Reimann, K. A., Reinhart, T. A., Rogan, M., Cavert, W., Miller, C. J., Veazey, R. S., Notermans, D., Little, S., Danner, S. A., Richman, D. D., Havlir, D., Wong, J., Jordan, H. L., Schacker, T. W., Racz, P., Tenner-Racz, K., Letvin, N. L., Wolinsky, S., and Haase, A. T. (1999). Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. *Science*, 286(5443):1353–1357.

Zheng, R., Jenkins, T. M., and Craigie, R. (1996). Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proceedings of the National Academy of Sciences*, 93(24):13659–13664.

Zychlinski, D., Schambach, A., Modlich, U., Maetzig, T., Meyer, J., Grassman, E., Mishra, A., and Baum, C. (2008). Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Molecular therapy : the journal of the American Society of Gene Therapy*, 16(4):718–725.

1

Curriculum Vitae

Personalia

Name	Lenard Sebastiaan VRANCKX
Nationality, gender	Belgian, male
Date & place of birth	February 17th 1989 - Leuven
Phone number	+32 (0) 478833700
E-mail	Lenard.vranckx@kuleuven.be Lenardvranckx@gmail.com (Personal)
Linguistic proficiency	Dutch - Native speaker English - fluently written and spoken French, Spanish - thorough knowledge
Current affiliation	Laboratory for Molecular Virology and Gene Therapy Department of Pharmaceutical and Pharmacological Sciences KU Leuven Kapucijnenvoer 33 VCTB+5 B7001 3000 Leuven, Belgium Phone number +32 (0) 16 374030 / Fax number +32 (0) 16 336336 http://www.kuleuven.be/molmed/
Research areas	Retrovirology Gene therapy Vectorology Next Generation Sequencing Molecular biology Biochemistry & Biotechnology
LinkedIn	https://be.linkedin.com/in/lenardvranckx

Education

2012-present *PhD in Biomedical Sciences*
KU Leuven, Faculty of Medicine
Department of Pharmaceutical and Pharmacological Sciences
DMolecular Medicine program

2011 *Dissertation in Virology and Gene Therapy*
Targeted integration of lentiviral vectors in the human genome -
LEDGF hybrids as tethering factors for lentiviral integration
Prof. Zeger Debyser and Prof. Rik Gijsbers
Laboratory for Molecular Virology and Gene Therapy
Department of Pharmaceutical and Pharmacological Sciences KU Leuven

2009-2011 *Master of Science (Magna Cum Laude)*
KU Leuven, Faculty of Sciences
Department of Chemistry
Biochemistry and Biotechnology program

2006-2009 *Bachelor of Science (Magna Cum Laude)*
KU Leuven, Faculty of Sciences
Department of Chemistry
Biochemistry and Biotechnology program

International Mobility

2010-2011 ERASMUS Exchange Programme
Universidad de Barcelona (E BARCELO01)

Working Experience

2012-now FWO Doctoral Fellow
Laboratory for Molecular Virology and Gene Therapy
Department of Pharmaceutical and Pharmacological Sciences
KU Leuven

Personal Grants, Scholarships

2014-2016 Doctoral Fellowship of The Fonds Wetenschappelijk Onderzoek (FWO) Renewal

2012-2014 Doctoral Fellowship of The Fonds Wetenschappelijk Onderzoek (FWO)

Courses and Training

9.9.2015 Usage of High Throughput Sequencing in Virology
Belgian Society for Virology, University of Liege, Liege, Belgium

1-4.7.2013 The Epigenetics Revolution: Development and Disease Biology Revisited

Teaching Experience

2016-present *Supervision junior PhD student Anne Bruggemans*
Biomedical Sciences program, KU Leuven
The Impact of LEDGIN-mediated retargeting of HIV integration on reactivation of latent HIV.
Prof. Zeger debyser
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

2015-present *Supervision junior PhD student Gerlinde Vansant*
Pharmaceutical Sciences program, KU Leuven
Study of the mechanism of action of a novel class of antivirals coined LEDGINs and their possible application in the quest for an HIV cure
Prof. Zeger debyser
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

2014-2015 *Co-supervision Master dissertation Simone Giovannozzi*
(ERASMUS exchange programme)
Molecular biology programme, University of Pisa, Italy
Retargeting of MLV integration using artificial p12 chimeras
Prof. Rik Gijsbers
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

Teaching Experience

August 2015 *Supervision Master Lab Rotation Student Jef Vanhamel*
Medicine program, KU Leuven
Prof. Zeger debyser
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

February 2015 *Supervision Master Lab Rotation Student Steve Ravelingien*
Medicine program, KU Leuven
Prof. Zeger debyser
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

August 2014 *Supervision Bachelor Lab Rotation Student Noemie Thomas*
Medicine program, KU Leuven
Prof. Zeger debyser
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

April 2014 *Supervision Master Lab Rotation Student Maria Loos*
Medicine program, KU Leuven
Prof. Zeger debyser
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium

2013-2014 *Co-supervision Bachelor Dissertation Sander Willox*
Optimization of lentiviral vectors with bioluminescent reporter technology
Laboratory for Molecular Virology and Gene Therapy, KU Leuven, Belgium
Prof. Rik Gijsbers
Biomedical laboratory technology program, KH Leuven, Belgium.

2013-present *Teaching Assistant Cell Biology Practical Course*
Prof. Zeger Debyser and Prof. Rik Gijsbers
Medicine and Biomedical Science programs, KULeuven, Belgium

2009-2010 *Tutor in Peer-Assisted Learning program*

Scientific Publications

Meeting abstracts - poster and oral presentations

Van Looveren, D., Elashkar, S., Schenk, F., Vranckx, L.S., De Rijck, J., Debyser, Z., MÃűdlich, U., Gijsbers, R. (2016). Artificial IN-peptide fusions provide BET-independent MLV-vectors with a safer integration profile. Human Gene Therapy: vol. 27. ESGCT. Florence, Italy, 18-21 October 2016, A182, Abstract No. P451.

Lampi, Y., Van Looveren, D., Vranckx, L.S., Bornschein, S., Debyser, Z., Gijsbers, R. (2016). Targeted editing of the PSIP1 gene encoding LEDGF/p75 as a functional cure strategy for HIV. Human Gene Therapy: vol. 27. ESGCT. Florence, Italy, 18-21 October 2016, A137, Abstract No. P300.

Debyser, Z., Vranckx, L.S., Demeulemeester, J., Saleh, S., Verdin, E., Cereseto A., Christ, F., Gijsbers, R. (2016). LEDGIN-mediated inhibition of the integrase-LEDGF/p75 interaction reduces reactivation of residual latent HIV. FRVC. Erlangen, Germany, 12-14 September 2016.

Vranckx, L.S., Demeulemeester, J., Saleh, S., Boll, A., Schrijvers, R., Battvelli, E., Verdin, E., Cereseto, A., Christ, F., Gijsbers, R., Debyser, Z. (2016). LEDGIN-mediated inhibition of the integrase-LEDGF/p75 interaction reduces reactivation of residual latent HIV. Retroviruses. Cold Spring Harbor, NY, USA, 23-28 May 2016.

Debyser, Z., Vranckx, L.S., Demeulemeester, J., Boll, A., Schrijvers, R., Verdin, E., Saleh, S., Christ, F., Gijsbers, R. (2016). Inhibition of the IN-LEDGF/p75 interplay by LEDGINs reduces reactivation from latency. CROI. Boston, MA, USA, 22-25 February 2016.

Vranckx, L.S., Demeulemeester, J., Boll, A., Schrijvers, R., Verdin, E., Cereseto, A., Saleh, S., Christ, F., Gijsbers, R., Debyser, Z. (2015). LEDGIN mediated inhibition of the IN-LEDGF/p75 interaction reduces reactivation from latency. Belvir Meeting. Brussels, Belgium, 18 December 2015.

Vranckx, L.S., Demeulemeester, J., Debyser, Z., Gijsbers, R. (2015). Safer, more randomized lentiviral vector integration using artificial LEDGF chimeras. ESGCT Congress. Helsinki, Finland, 17-20 September 2015.

Boll, A., Vranckx, L.S., Gijsbers, R., Debyser, Z., Cereseto, A. (2013). LEDGF/p75 controls 3D-localization of HIV-1 provirus in the nuclear compartment of infected cells. Frontiers of Retrovirology. Cambridge, UK, 16-18 September 2013.

Journals with peer review

Vranckx, L.S., Demeulemeester, J., Debyser, Z., Gijsbers, R. (2016). Towards a safer, more randomized lentiviral vector integration profile exploring artificial LEDGF chimeras. PLoS One, 11(10), art.nr. 10.1371/journal.pone.0164167.

Vranckx, L.S., Demeulemeester, J., Saleh, S., Boll, A., Vansant, G., Schrijvers, R., Weydert, C., Battivelli, E., Verdin, E., Cereseto, A., Christ, F., Gijsbers, R., Debyser, Z. (2016). LEDGIN-mediated inhibition of integrase-LEDGF/p75 interaction reduces reactivation of residual latent HIV. EBioMedicine, Volume 8 , 248 - 264.

Quercioli, V., Primio, C., Casini, A., Mulder, L., Vranckx, L.S., Borrenberghs, D., Gijsbers, R., Debyser, Z., Cereseto, A. (2016). Comparative analysis of HIV-1 and MLV 3D nuclear distribution. Journal of Virology, 90 (10), art.nr. JVI.03188-15, 5205-5209.

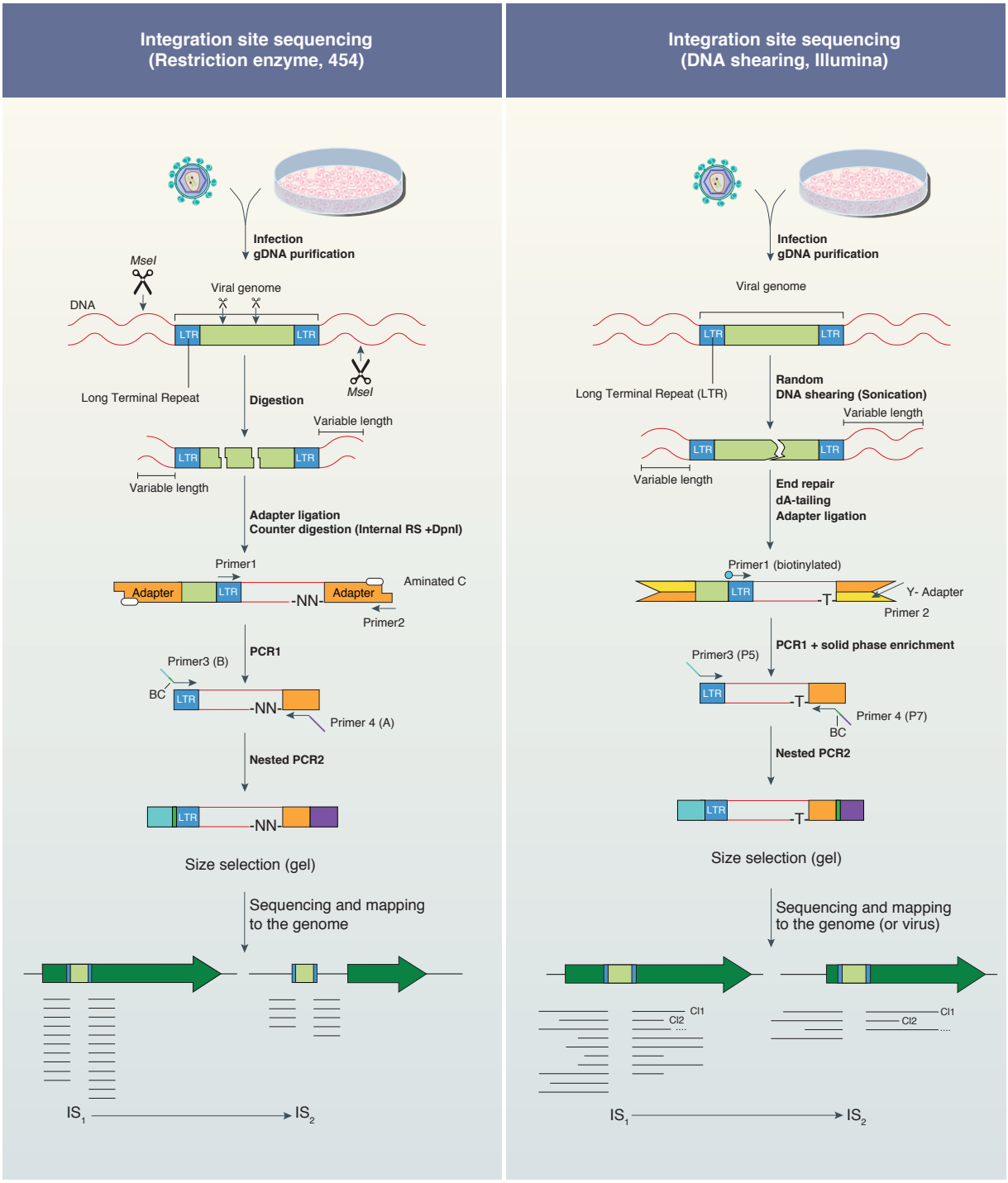
Articles in review or preparation

Lévy, C., Amirache, F., Girard-Gagnepain, A., Frecha, C., Rodriguez, F.J.R., Bernadin, O., Costa, C., Nègre, D., Vranckx, L.S., Taylor, N., Clerc, I., Bueren, J., Rio, P., Gijsbers, R., Cosset, F., Verhoeyen, E. (2016). Measles virus envelope pseudotyped lentivectors transduce resting human hematopoietic stem cells at efficiency without precedent. In Submission

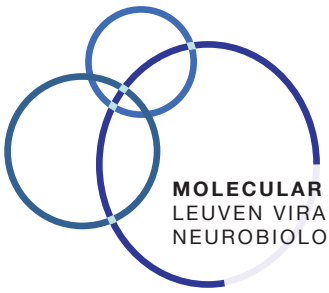
El Ashkar, S., Van Looveren, D., Schenk, F., Vranckx, L.S., Demeulemeester, J., de Rijck, J., Debyser, Z., Modlich, U., Gijsbers, R. (2016). Engineering Next Generation BET-independent MLV-Vectors for Safer Gene Therapy. In Submission

Saleh, S., Vranckx, L.S, Gijsbers, R., Christ, F., Debyser, Z. (2016) Insight in HIV-2 latency may disclose strategies for a cure of HIV-1 infection. (Review) In Submission

Appendix



Supplementary Figure A.1: Sequencing of proviral integration sites. Outlines for different integration site sequencing approaches.



MOLECULAR VIROLOGY & GENE THERAPY
LEUVEN VIRAL VECTOR CORE - LVVC
NEUROBIOLOGY & GENE THERAPY

KU Leuven
Pharmaceutical and Pharmacological Sciences
Molecular Virology and Gene Therapy
Kapucijnenvoer 33, VCTB +5 B7001
B-3000, Leuven